DOCUMENT RESUME

ED 130 042 CE 008 247

| | |
|---|---|
| AUTHOR | Dunnette, Marvin D.; Motowidlo, Stephan J. |
| TITLE | Police Selection and Career Assessment. |
| INSTITUTION | Personnel Decisions, Inc., Minneapolis, Minn. |
| SPONS AGENCY | National Inst. of Law Enforcement and Criminal Justice (Dept. of Justice/LEAA), Washington, D.C. |
| PUB DATE | Nov 76 |
| GRANT | 73-N1-99-0018-G; 74-N1-99-0001-G |
| NOTE | 208p. |
| AVAILABLE FROM | Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 20402 (Stock Number 027-000-00390-7, $2.85) |
| EDRS PRICE | MF-$0.83 HC-$11.37 Plus Postage. |
| DESCRIPTORS | Aptitude Tests; Employment Qualifications; *Evaluation Methods; Job Placement; *Occupational Tests; Performance Tests; *Personnel Evaluation; Personnel Management; *Personnel Selection; *Police; Promotion (Occupational); Psychological Evaluation; Simulation; Test Construction; Test Reviews; Test Validity |
| IDENTIFIERS | *Police Career Index |

ABSTRACT

Research was conducted to develop the Police Career Index (PCI) and the regional assessment center exercises to provide a total personnel evaluation system to help police departments screen applicants, evaluate on-the-job performance of officers eligible for promotion, and gauge a person's suitability for police work. The PCI is based on actual performance criteria developed through cooperation with police officials in nine cities throughout the country. It measures how well an individual might handle four different police jobs--patrol officer, patrol sergeant, investigator-detective, and mid-level commander. This easily administered, objectively scored written test screens out high-risk applicants. Those who register borderline scores are referred to a regional assessment center where a series of police job simulations elicit behavioral indicators of potential for success in various aspects of police work. Local departments may wish to consider the 2-fold system, tested with good results in several cities. About half of this report consists of description of the research to develop and validate the PCI and the assessment center exercises. Appendixes include job description booklets for precinct patrolman, investigative personnel, sergeants, and intermediate commanders (about 50 pages). The remainder is research data tables. (TA)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED130042

# POLICE SELECTION AND CAREER ASSESSMENT

By

Marvin D. Dunnette

and

Stephan J. Motowidlo

With the help of:

Robert W. Heckman, Andrea P. Hunt, Leaetta M. Hough, Rodney L. Rosse, David A. Bownas, Irene E. Roban, Paul D. Johnson, Robert Flint, Walter C. Borman, Lowell W. Hellervik, Dennis M. Groner, & Marlys Gimble.

This project was supported by Grant Numbers 73-N1-99-0018-G and 74-N1-99-0001-G awarded to Personnel Decisions, Inc., by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

NOVEMBER 1976

National Institute of Law Enforcement and Criminal Justice
Law Enforcement Assistance Administration
U.S. Department of Justice

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

CE 008 347

**NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE**

Gerald M. Caplan, *Director*

**LAW ENFORCEMENT ASSISTANCE ADMINISTRATION**

Richard W. Velde, *Administrator*

Henry F. McQuade, *Deputy Administrator*

Paul K. Wormeli, *Deputy Administrator*

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402—Price $2.85 cents
Stock Number 027-000-00390-7

# CONTENTS


| | Page |
|---|---|
| **FOREWORD** | xi |
| **ACKNOWLEDGEMENTS** | xii |
| **EXECUTIVE SUMMARY** | xv |
| A. Project Objectives | xv |
| B. Project Tasks and Methods | xv |
| C. Project Findings and Project Products | xvi |
| D. Combined Operational Use of Police Career Index and Regional Assessment Centers for Personnel Decisions and Career Guidance of Police Personnel | xviii |
| **CHAPTER I. BACKGROUND, RESEARCH NEEDS, AND RESEARCH APPROACH** | 1 |
| A. Background | 1 |
| B. Research Needs, Early Results, and Research Approach | 1 |
| **CHAPTER II. DEVELOPMENT OF THE POLICE CAREER INDEX (PCI)** | 4 |
| A. Sequence of Development | 4 |
| B. Selection and Development of Experimental Tests and Inventories | 4 |
| C. Assignments for Test Item Selection and Writing | 5 |
| D. Contents of Experimental Batteries | 7 |
| 1. Personal history/biographical inventory | 7 |
| 2. Opinion and self-description inventories | 7 |
| E. Administration of Tests and Collection of Performance Ratings | 8 |
| 1. Test administration | 8 |
| 2. Performance ratings | 10 |
| F. Analysis of Performance Ratings | 12 |
| 1. Overview of analysis and findings | 12 |
| 2. Reliability studies | 12 |
| 3. Factor analytic studies | 22 |
| G. Item Analysis and Development of Predictor Scales | 26 |
| 1. General overview of analyses | 26 |
| 2. Item analysis and cross-validation | 27 |
| 3. Modal response keys | 30 |
| 4. Factor analysis of predictor-criterion composites | 30 |
| 5. Item pools and criterion limits for item analysis | 30 |
| H. Rating, Predictor Scale, and Validity Information According to Cities and Subgroups | 41 |
| 1. Performance ratings | 41 |
| 2. Validities by subgroup | 44 |
| 3. Predictor scale scores by subgroup | 48 |
| I. Standard Scoring and Interpretive Information for POLICE CAREER INDEX | 50 |
| 1. Patrol Officer CAREER INDEX | 50 |
| 2. Detective CAREER INDEX | 51 |
| 3. Sergeant CAREER INDEX | 53 |
| 4. Middle Level Command Officer CAREER INDEX | 53 |
| 5. Operational Use of POLICE CAREER INDEX | 54 |

Page

**CHAPTER III. DEVELOPMENT OF POLICE ASSESSMENT CENTER EXERCISES** ..... 56

A. History of the Use of Assessment Centers ..... 56
B. Development of the Assessment Dimensions ..... 58
C. Development of the Simulation Exercises ..... 59
D. Catalogue of Simulation Exercises ..... 63
E. Pilot Tests of Police Assessment Centers ..... 66
F. Preparation of Assessors' Manuals ..... 75
G. Recommendations for Implementing Police Assessment Centers ..... 76

**CHAPTER IV. OPERATIONAL USE OF POLICE CAREER INDEX AND REGIONAL ASSESSMENT CENTERS FOR SELECTION AND CAREER GUIDANCE OF POLICE OFFICERS** ..... 78

**REFERENCES** ..... 81

**APPENDIX A. Job Performance Description Booklets For:**

Precinct Patrolmen ..... 84
Investigative Personnel ..... 101
Sergeants ..... 119
Intermediate Commanders ..... 133

**APPENDIX B. Tables Showing Interrater Reliability Estimates for Job Performance Ratings Separately, by City, for Detectives, Sergeants, and Middle Command Personnel** ..... 149

**APPENDIX C. Sample Sizes, Means and Standard Deviations of Job Performance Dimension Ratings; and Correlation Matrices Between Performance Ratings for Each of Four Police Job Areas** ..... 161

**APPENDIX D. Decision Rules Used for Developing Modal Response Keys for Inventories of Likes and Dislikes and Situational Judgment Inventories** ..... 167

**APPENDIX E. Tables Showing Numbers of Weighted Items, Median Foldback and Cross-Validities for Various Scoring Keys According to Item Pool and Criterion Score Limits for Each Criterion Dimension and Each Police Job Function** .... 171

**APPENDIX F. Technical Note: Procedure Used for Estimating Expected Cross-Validity Coefficients for Final Predictor Scales.** 181

**APPENDIX G. Item Samples and Illustrative Scoring Weights for Career Index Predictor Scales for Patrol Officers, Detectives, Sergeants, Middle Level Command Officers** ..... 185

**APPENDIX H. Validities for Predictor Scales Against Criterion Dimensions for Total Samples and Separately by City** .... 193

**APPENDIX I. Police Career Index Booklets***

**APPENDIX J. Assessor's Manuals***

*Materials comprising these appendixes are bulky and in short supply. Information about obtaining copies of these appendixes may be obtained by writing to Dr. Marvin D. Dunnette, Personnel Decisions, Inc., 821 Marquette Avenue, Foshay Tower, Minneapolis, Minnesota 55402.

# TABLES


| Number | | Page |
|---|---|---|
| 1 | Data Collection Sequence According to Time, City, and Type of Police Jobs | 9 |
| 2 | Numbers of Police Officers Tested and Rated According to City and Type of Police Job | 9 |
| 3 | Numbers of Patrol Officers Whose Job Performance Was Rated by 1, 2, 3, Etc. Raters | 13 |
| 4 | Interrater Reliability Estimates for Job Performance Ratings for all Patrol Officers (N = 886) | 13 |
| 5 | Interrater Reliability Estimates for Job Performance Ratings for Cincinnati Patrol Officers (N = 142) | 14 |
| 6 | Interrater Reliability Estimates for Job Performance Ratings for Des Moines Patrol Officers (N = 114) | 14 |
| 7 | Interrater Reliability Estimates for Job Performance Ratings for Miami (Dade County) Patrol Officers (N = 98) | 15 |
| 8 | Interrater Reliability Estimates for Job Performance Ratings for Minneapolis Patrol Officers (N = 238) | 15 |
| 9 | Interrater Reliability Estimates for Job Performance Ratings for Portland Patrol Officers (N = 130) | 16 |
| 10 | Interrater Reliability Estimates for Job Performance Ratings for Washington Patrol Officers (N = 164) | 16 |
| 11 | Interrater Reliability Estimates for Retained Job Performance Ratings for All Tested Patrol Officers | 17 |
| 12 | Numbers of Detectives Whose Job Performance Was Rated by 1, 2, 3, & 4 Raters | 18 |
| 13 | Interrater Reliability Estimates for Job Performance Ratings for all Detectives (N = 483) | 18 |
| 14 | Interrater Reliability Estimates for Retained Job Performance Ratings for all Tested Detectives | 19 |
| 15 | Numbers of Sergeants Whose Job Performance Was Rated by 1, 2, 3, Etc. Raters | 19 |
| 16 | Interrater Reliability Estimates for Job Performance Ratings For All Sergeants (N = 379) | 20 |
| 17 | Interrater Reliability Estimates for Retained Job Performance Ratings for all Tested Sergeants | 20 |
| 18 | Interrater Reliability Estimates for Job Performance Ratings for all Middle Level Officers (N = 215) | 21 |
| 19 | Interrater Reliability Estimates for Retained Job Performance Ratings for all Tested Middle Level Officers | 21 |

| Number | | Page |
|---|---|---|
| 20 | Median Reliability Estimates, Dimensions With Highest Reliabilities, and Reliabilities for Overall Rating and Mean of Ratings for Retained Performance Ratings | 22 |
| 21 | Factor Matrices for Two-Factor and for Three-Factor Solutions for Job Performance Ratings for Patrol Officers | 23 |
| 22 | Factor Matrices for Three-Factor and for Four-Factor Solutions for Job Performance Ratings for Detectives | 24 |
| 23 | Factor Matrix Showing Three-Factor Solution for Job Performance Ratings for Sergeants | 25 |
| 24 | Factor Matrix Showing Four-Factor Solution for Job Performance Ratings for Middle Level Command Officers | 25 |
| 25 | Dimensions Chosen for Item Analyses; Reliabilities; and Numbers of Tested Officers for Whom Ratings Were Available | 26 |
| 26 | Correlation Coefficients Between LD and SITNL Modal Response Keys and Criterion Dimension Ratings for Patrol Officers | 31 |
| 27 | Initial Scoring Keys Chosen for Patrol Officer Criterion Dimensions: Criterion Score Limits, Numbers of Items, Foldback Coefficients, and Median Cross-Validities | 32 |
| 28 | Scoring Keys for Patrol Officer Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against P-Random | 33 |
| 29 | Factor Matrix for Four-Factor Solution of Scoring Key-Criterion Dimension Matrix for Patrol Officers | 33 |
| 30 | Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Four Scoring Keys (Predictor Scales) for the Patrol Officer Career Index | 34 |
| 31 | Intercorrelations Between Predictor Scales Comprising the Patrol Officer Career Index | 34 |
| 32 | Correlation Coefficients Between LD And SITNL Modal Response Keys and Criterion Dimension Ratings for Detectives | 34 |
| 33 | Scoring Keys for Detective Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against D-Random | 35 |
| 34 | Factor Matrix for Three-Factor Solution of Scoring Key-Criterion Dimension Matrix for Detectives | 35 |
| 35 | Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Two Scoring Keys (Predictor Scales) for the Detective Career Index | 36 |
| 36 | Correlation Coefficients Between LD and SITNL Modal Response Keys and Criterion Dimension Ratings for Sergeants | 36 |
| 37 | Scoring Keys for Sergeant Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against S-Random | 36 |
| 38 | Factor Matrix for Two-Factor Solution of Scoring Key-Criterion Dimension Matrix for Sergeants | 37 |

| Number | | Page |
|---|---|---|
| 39 | Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Two Scoring Keys (Predictor Scales) for the Sergeant Career Index | 37 |
| 40 | Correlation Coefficients Between LD and SITNL Modal Response Keys and Criterion Dimension Ratings for Middle Level Command Officers | 38 |
| 41 | Scoring Keys for Middle Command Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against C-Random | 38 |
| 42 | Factor Matrix for Four-Factor Solution of Scoring Key-Criterion Dimension Matrix for Middle Level Command Officers | 39 |
| 43 | Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Five Scoring Keys (Predictor Scales) for the Middle Level Command Career Index | 39 |
| 44 | Intercorrelations Between Predictor Scales Comprising the Middle Command Officer Career Index | 40 |
| 45 | Item Content, Median Reliabilities, Median Correlations Against Criterion and Pseudo-Criterion (Random) Distributions, Median Estimated Cross-Validities, and Median Intercorrelations Between Predictor Scales Comprising Four Career Index Inventories | 40 |
| 46 | Means and Standard Deviations on Criterion Dimensions and T Values for Minneapolis and Portland Patrol Officers Who Were Tested and Not Tested | 42 |
| 47 | Means and Standard Deviations of Ratings on Overall Criterion Dimensions According to City and Police Functions (White Males Only) | 42 |
| 48 | Numbers of Men, Women, Whites and Non-Whites According to Police Function | 42 |
| 49 | Means and Standard Deviations on Criterion Dimensions for White and Non-White Patrol Officers | 43 |
| 50 | Means and Standard Deviations on Criterion Dimensions for Male, Female, White and Non-White Detectives | 43 |
| 51 | Means and Standard Deviations on Criterion Dimensions for White and Non-White Sergeants | 43 |
| 52 | Means and Standard Deviations on Criterion Dimensions for White Middle Level Command Officers | 43 |
| 53 | Means and Standard Deviations on Overall Performance Ratings for Officers Above and Below the "Midpoint" on Height, Weight, Age, and Length of Police Service (White Males Only) | 45 |
| 54 | Range of Values and Median Values for Correlations Between Cognitive Tests and Criterion Dimension Ratings and Between Cognitive Tests and Predictor Scales of the Career Index | 45 |
| 55 | Mean Absolute Deviations Between Total Sample Validities and Validities for Separate Cities According to Predictor Scale | 46 |
| 56 | Validities of Patrol Officer Predictor Scales Against all Criterion Dimensions for Total Sample and for Minority Officers | 47 |

| Number | | Page |
|---|---|---|
| 57 | Validities of Detective Predictor Scales Against all Criterion Dimensions for Total Sample, Minority and Female Officers | 47 |
| 58 | Validities of Sergeant Predictor Scales Against all Criterion Dimensions for Total Sample and for Minority Officers | 48 |
| 59 | Means and Standard Deviations on Patrol Officer Career Index Predictor Scales According to Race of Examinee | 49 |
| 60 | Means and Standard Deviations on Detective Career Index Predictor Scales According to Race of Examinee | 49 |
| 61 | Means and Standard Deviations on Sergeant Career Index Predictor Scales According to Race of Examinee | 49 |
| 62 | Distribution of Police Officers Scheduled to Participate in the Pretesting of the Simulation Exercises, According to City and Police Function | 62 |
| 63 | Assessment Dimensions Taped by the Patrol Officer Exercises | 67 |
| 64 | Assessment Dimensions Taped by the Detective Exercises | 68 |
| 65 | Assessment Dimensions Taped by the Sergeant Exercises | 68 |
| 66 | Assessment Dimensions Taped by the Intermediate Commander Exercises | 69 |
| 67 | Cities in Which the Pilot Assessment Centers Were Conducted | 69 |
| 68 | Schedule for Assessors and Candidates in the Pilot Patrol Officer Center | 70 |
| 69 | Schedule for Assessor I in the Pilot Detective Center | 70 |
| 70 | Schedule for Assessor II in the Pilot Detective Center | 71 |
| 71 | Schedule for Assessor III in the Pilot Detective Center | 71 |
| 72 | Schedule for Candidate A in the Pilot Detective Center | 72 |
| 73 | Schedule for Candidate B in the Pilot Detective Center | 72 |
| 74 | Schedule for Candidate C in the Pilot Detective Center | 73 |
| 75 | Schedule for Assessors and Candidates in the Sergeant Center | 73 |
| 76 | Schedule for Assessors and Candidates in the Intermediate Commander Center | 74 |

# FIGURES


| Number | | Page |
|---|---|---|
| 1 | Stages of Personnel Selection, Performance Evaluation, Assessment, and Career Placement Decisions | xix, 3 |
| 2 | Rating Worksheet for Sergeants' Job Category A: CONCERN FOR SUBORDINATES | 11 |
| 3 | Sample Report Format for Results from the Patrol Officer POLICE CAREER INDEX | 51 |
| 4 | Sample Report Format for Results from the Detective POLICE CAREER INDEX | 52 |
| 5 | Sample Report Format for Results from the Sergeant POLICE CAREER INDEX | 53 |
| 6 | Sample Report Format for Results from the Middle Level Command Officer POLICE CAREER INDEX | 55 |
| 7 | Stages of Personnel Selection, Performance Evaluation, Assessment, and Career Placement Decisions | 79 |

# FOREWORD

Selecting the right person for the job is never easy. For police departments, it's particularly difficult. Enforcing the law is a complex, sensitive task that makes great demands on the individual officer.

Because the stakes are so high, police departments have turned increasingly to a variety of tests to help them screen applicants and evaluate on-the-job performance of officers eligible for promotion. At the same time, police personnel experts have emphasized the need to develop more reliable methods for gauging a person's suitability for police work. One new approach developed and tested with Institute support is described in this report.

The research produced a Police Career Index based on actual performance criteria that measures how well an individual might handle four different police jobs—patrol officer, patrol sergeant, investigator-detective, and mid-level commander. This easily administered written test screens out high-risk applicants. Those who register borderline scores are referred to a Regional Assessment Center for rigorous psychological evaluation.

The two-fold procedure was tested with good results in several cities. Local departments may wish to consider the proposed system for use in improving their personnel decisions.

Gerald M. Caplan,
Director
National Institute of Law Enforcement
and Criminal Justice

# ACKNOWLEDGMENTS

Many persons have helped in this research over the many years it has been under way.

We are grateful to the several officials from the Law Enforcement Assistance Administration in Washington, D. C., who have served in the role of Grant Monitor to us during various phases of the research program. These persons include Mr. Sidney Epstein, Dr. Richard Laymon, Mr. Lewis Mayo, and Ms. Kay Monte.

We also appreciate the excellent assistance we received from officials working in the nine police departments which cooperated with us in many different ways during the gathering of data for developing the POLICE CAREER INDEX. These persons include Captain Myron Blanch of Minneapolis, Minnesota; Mr. Harvey McGowan and Captain Robert Schwartz of Portland, Oregon; Mr. John Palmer and Captain Norman Hughes of Cincinnati, Ohio; Sergeant F. Carson and Mr. R. V. Sipfle of Des Moines, Iowa; Lieutenant Charles Ellison and Captain Larry Gore of San Diego, California; Mr. Donald Slesnick, Ms. Faye Trimble, and Ms. Mary Jean Fitzgibbons of Miami (Dade County), Florida; Dr. Brian O'Leary of Washington, D. C.; Sergeant Richard Milne of Tucson, Arizona; and Sergeant Rivera of Albuquerque, New Mexico.

Many persons helped us during all phases of the development, debugging, and pretesting of the simulation and assessment center exercises. We thank them for their time and their insightful suggestions. Included among these persons are Captain Allan Hoehl of New York City, New York; Dr. Robert Mills of Cincinnati, Ohio; Police Chief Donald Derning of Winnetka, Illinois; Captain Myron Blanch of Minneapolis, Minnesota; David Gorski of Golden Valley, Minnesota; Dr. Michael Roberts of San Jose, California; Captain Robert Schwartz of Portland, Oregon; Dr. Terry Eisenberg of San Jose, California; Dr. Brian O'Leary of Washington, D. C.; Mr. Pierce Brooks of Lakewood, Colorado; Dr. Martin Reiser of Los Angeles, California; and Captain James Shaw of King County, Washington.

In addition to our own staff members, many persons worked with us as staff members during our pretests of the various police assessment centers. Obviously, their help was crucial to the success of this phase of the research. Those persons include Dr. Milton Davis, Assistant Chief Ronald Still, Dr. Paul Reiling, Dr. Chad Karr, Commander James Brouillette, and Mr. Pat Fleming, all of Portland, Oregon; Captain Calvin Nelson, Mr. James Stengel, and Mr. Paul Maloney, all of Minneapolis, Minnesota; Ms. Sheila Thompson, Dr. Walter Mann, and Captain Theodore Carr, all of Washington, D. C.; and Dr. Willard Lewis, Mr. Robert Overman, Detective Chief Donald Miller, Chief Jerry Pitts, Dr. Richard Arvey, and Reverend Ted Baker, all of Chattanooga, Tennessee.

We also are deeply in debt to the over 2,000 police officers from all over the country who served so willingly and helpfully as subjects (''guinea pigs'') for the many phases of our research.

Perhaps our biggest debt of gratitude is owed to the often unsung, but exceptionally gallant and hard working persons who manage to convert masses of nearly illegible numbers, tables, handwriting, misspellings, and poor syntax into a finished product that is exceptionally pleasant to look at and even readable. For us, these wonderful persons include Ms. Kathy Moe, Ms. Jana Odegaard, Ms. Elin Gowdy, and Ms. Becky Payne. We owe them much. The simplicity and attractiveness of the cover of this report was the work of Ms. Mary Towner. We appreciate her artistic flair.

Ms. Marlys Gimble has been responsible over these many years for managing to keep track of schedules, budgets, coordinating printing and duplicating efforts, and just generally helping us maintain a semblance of sanity when there seemed little rational reason for doing so.

Finally, we thank all our other staff members who are listed on our cover page. All played strong supporting roles in the several phases of this research. The contributions of each was unique, ranging from planning the research, securing design and securing cooperation from cities or administering tests, to "cleaning up" the incoming data or making certain of the wisdom of statistical procedures used in the analyses.

We sincerely hope that our own efforts in summarizing all these results may reflect, to a degree at least, the high standards and conscientious efforts put forth by all these people.

Marvin D. Dunnette
Stephan J. Motowidlo

# EXECUTIVE SUMMARY

## A. Project Objectives

Our major objective in this research program has been to develop new methods for evaluating persons who apply for positions in police work and for assessing the potential of present police officers who are being considered for promotion.

In order to accomplish this central objective, an important secondary objective was to learn as fully as possible exactly what activities or job behaviors are critical to effectiveness and ineffectiveness (success or failure) in police work. Thus, a first step in the research program involved studying the critical features of four different police jobs: general patrol officer, patrol sergeant, investigator/detective, and the intermediate command function. Defining these critical dimensions of job performance for each of these four police functions was used in two distinct ways to accomplish the central objective of developing new assessment methods for police jobs:

- First, the critical dimensions of police performance were used to design new methods of rating the job performance of police officers in the four police functions mentioned above. These rating methods were used to evaluate the current job performance effectiveness of police officers in order to conduct a concurrent validation study toward the end of constructing a short, easily administered inventory with separate scoring keys shown empirically to be related to different levels of performance effectiveness in those four functions.
- Second, the critical dimensions of police performance were the basis for designing simulations and standardized situational tasks for use in evaluating candidates being considered for placement in any of the four police functions. In effect, these "work samples" or job simulations were designed to be used in assessment centers where the success potential of candidates for various police jobs can be evaluated behaviorally.

In essence, then, the central objective of this research was to develop two quite different but complementary procedures for evaluating candidates for any of four police jobs. One, which came to be called the POLICE CAREER INDEX, consists of a brief, easily administered and objectively scored inventory to be used in preliminary screening of applicants and candidates for police jobs. The other procedure, the Regional Assessment Center, consists of a series of police job simulations tailor-made to elicit behavioral indicators of a person's potential for success in various aspects of police work. The intention is that these assessment centers will be available on a regional basis for use in evaluating career potential, strengths and weaknesses, and training or development needs for police job candidates and for currently employed police officers.

## B. Project Tasks and Methods

Development of the POLICE CAREER INDEX was accomplished with the cooperation of police officials in nine cities located throughout the country. Job performance ratings were obtained for police officers from each city. In addition, most of the persons who were rated also completed a comprehensive experimental battery of tests and inventories. Police officers participating in these activities did so, for the most part, on off-duty time, and they were reimbursed for their time. The numbers of officers for whom tests and ratings were finally available are shown below according to city:

| | *Completed Tests and Inventories* | *Job Performance Ratings Available* |
|---|---|---|
| Albuquerque | 18 | 18 |
| Cincinnati | 294 | 315 |
| Des Moines | 201 | 218 |
| Miami (Dade Co.) | 239 | 277 |
| Minneapolis | 309 | 360 |
| Portland, Oregon | 169 | 244 |
| San Diego | 140 | 140 |
| Tucson | 28 | 28 |
| Washington, D. C. | 419 | 404 |
| Totals | 1,817 | 2,004 |

After discarding materials which were incomplete or improperly filled in, complete responses and performance ratings were available for 700 patrol officers, 415 detectives, 362 sergeants, and 206 middle command officers (lieutenants and captains).

A variety of statistical methods of analysis was used to examine the dimensionality of the performance ratings and to identify patterns of responses to test and inventory items which were most highly correlated with various performance ratings. A Monte Carlo method of item analysis and scoring key development was used to select scoring keys most likely to yield stable validity coefficients. These methods were applied separately to each of the four police functions in developing POLICE CAREER INDEX keys. After scoring keys had been developed, their validities were examined separately according to subgroupings based on city, race, and sex.

Development of simulations to be used in the Regional Assessment Center program was carried out over a series of steps directed toward continual examination, modifi-

cation, and finer and finer tuning of the standardized exercises to be used in each of the four police assessment centers. These steps are listed and briefly described in the following:

- Intensive study of critical performance dimensions for each police function was supplemented with firsthand observations by our staff members of officers as they carried out their jobs. Staff members developed a total of 55 preliminary job simulation ideas which they believed sampled the major dimensions across the four police functions.
- These preliminary simulations were sent to nine leading police psychologists and police officials who provided critical reviews and feedback concerning their appropriateness for use in assessing police job candidates.
- Based on the above feedback, the most promising seven to ten simulations for each function were elaborated to give them realism. Instructions for administration were written and necessary physical props were provided to round out as fully as possible each of the exercises.
- The above sets of police job simulation exercises were then critiqued in detail during an intensive one-day working conference of 11 police psychologists and police officials. Attendees came from departments all over the country to attend the conference in Minneapolis. As a result, some exercises were eliminated, important modifications were made in all of them, and new simulations were designed for some of the functions. Output from this conference yielded a total of 30 fully elaborated job simulations: eight for patrol officers, eight for detectives, five for sergeants, and nine for intermediate commanders.
- The simulations were then pretested by actually administering them to police officers in the three cities of Minneapolis, Des Moines, and Cincinnati. A total of 77 officers agreed to participate in these pretests, including 16 patrol officers, 22 detectives, 25 sergeants, and 14 intermediate commanders. Each of the 30 simulations was pretested on an average of six police officers. Their comments and reactions during and after the pretests were rich in suggestions for final modifications to improve and "finalize" each of the simulations. At this stage, four more exercises were dropped leaving a total of 26: 7 each for patrol officers, sergeants, and middle commanders, and 5 for detectives.
- In order to provide a final full-scale evaluation of the assessment exercises, eight pilot assessment centers were held as shown below:

*Pilot Assessment Centers*

| City | Type | | | |
|---|---|---|---|---|
| | Patrol | Detective | Sergeant | Command |
| Chattanooga | | | X | X |
| Minneapolis | X | X | | |
| Portland, Oregon | | | X | X |
| Washington, D. C. | X | X | | |

These final pretests were highly successful. Even with minimal training, assessor staff members were able to make sound behavioral observations, perform accurate ratings of candidates, and come to good agreement about the relative overall effectiveness of the candidates whom they observed.

## C. Project Findings and Project Products

1. *Scoring keys.* Our analyses of the statistical relationship between job performance ratings and responses to items in the experimental test battery yielded several scoring keys for each of the four police functions. These keys are listed below for each of the four functions:

*Patrol officer keys*

- Key PI, called *Public Contact,* is most closely related to those aspects of the officer's job involving dealing constructively and effectively with the public on a one-to-one basis.
- Key PII, called *Overall Performance,* is most closely related to ratings of overall job effectiveness as a patrol officer.
- Key PIII, called *Crime Prevention,* is most closely related to those aspects of the officer's job involving alertness to suspicious situations, detecting and investigating crimes, and maintenance of public safety.
- Key PIV, called *Cooperation,* is most closely related to those aspects of the officer's job involving effective cooperation with other officers and other law enforcement units.

Our analyses suggest that the minimum estimated median validity for these patrol officer keys is .24. When scores from them are combined to form a composite TOTAL POTENTIAL score, the validity may be in the range .30 to .35.

*Detective keys*

- Key DI, called *Overall Performance,* is most closely related to ratings of overall job effectiveness as a detective.
- Key DII, called *Personal Integrity,* is related to those aspects of a detective's job involving fairness and integrity in dealing with all cases equally and resisting opportunities to use one's badge for personal gain.

Our analyses suggest that the minimum estimated median validity for these detective keys is .25. When scores from them are combined to form a composite TOTAL POTENTIAL score, the validity may be in the range of .35.

*Sergeant keys*

- Key SI, called *Supervisory Coordination*, is related to those aspects of a sergeant's job involving inspection of personnel and equipment, scheduling duty time, and deploying officers and equipment with wisdom and initiative.
- Key SII, called *Supervisory Consideration*, is related to those aspects of a sergeant's job involving awareness of subordinates' needs, recognizing and praising good performance, and training or orienting subordinates toward good performance.

Our analyses suggest that the minimum estimated median validity for these sergeant keys is .20. When scores from them are combined to form a composite TOTAL POTENTIAL score, the validity may be in the range of .25.

*Middle level command keys*

- Key CI, called *Administration*, is related to those aspects of a middle commander's job involving administrative and scheduling duties, such as paperwork, assigning manpower, and organizing office procedures.
- Keys CII and CIII both are called *Overall Performance*, and both are related most closely to ratings of overall job effectiveness in middle level command positions (lieutenants and captains).
- Key CIV, called *Dedication*, is related to those aspects of a middle commander's job involving dedication, setting a good example, and resisting opportunities to use one's position for personal gain.
- Key CV, called *Personnel*, is related to those aspects of a middle commander's job involving working with subordinates in areas related to their evaluation, motivation, discipline, and development.

Our analyses suggest that the minimum estimated median validity for these middle level command keys is .30. When scores from them are combined to form a composite TOTAL POTENTIAL score, the validity may be in the range of .35 to .40.

2. *Validities by city, by race, and by sex.* City-by-city comparison of the validities of the above scoring keys shows that they are impressively consistent and that they are very similar to those obtained for the overall samples. Not only are the actual magnitudes of validity coefficients similar from city to city, but the patterns of higher and lower values are similar over all cities. These results suggest that the scoring keys may be used as they stand to help make personnel decisions about applicants for police work and candidates for promotion. However, such use should, of course, be accompanied by studies to establish local norms and local validity information specific to each department.

A sufficient number of female officers was available for separate validity analysis only in detective jobs. Validities for the small group of female detectives (numbering only 15 to 25) were lower for the DI key but about the same for the DII scale as for all detectives combined.

Sufficient numbers of minority officers were available for separate validity analysis for patrol officer, detective, and sergeant jobs. Validities are essentially the same for whites and for non-whites for the patrol officer and sergeant scoring keys. In contrast, validities for the detective scoring keys are uniformly lower for minority officers than for officers in the total sample of detectives. Studies of mean differences in performance ratings and in scoring key scores between white and non-white officers showed that adjustments needed to be made for minority candidates in order to assure that persons, regardless of race, with equal probabilities of showing particular levels of job performance effectiveness will obtain approximately the same scores on the TOTAL POTENTIAL composites of the POLICE CAREER INDEX scoring systems.

At present, we can recommend use of these scoring systems for both minority and non-minority candidates for patrol officer and sergeant police functions. Use of the scoring systems for women and for minority persons for the detective function is not currently recommended. Evidence concerning possible sex differences is not available from our research for police functions of patrol officer, sergeant, and middle command. Evidence about possible race differences is not available from our research for middle command police functions.

3. *Operational use of POLICE CAREER INDEX.* The POLICE CAREER INDEX consists of three inventories. The first inventory contains 393 items made up of background information, personality-type items, and items involving preferences for various activities. It contains all material necessary for scoring the patrol officer and detective keys and most of the sergeant and middle command keys. The second inventory contains 82 job situation items relevant specifically to the sergeant function, and the third inventory contains 75 job situation items relevant specifically to the middle command function. The inventories can be administered quickly and easily by a competent clerical person to groups of candidates numbering as large as 75 to 100. The time required for administration is no more than 2 hours (the average

time for a candidate is about 75 minutes) for the Biographical and Personal Information Inventory and about 45 minutes for either one of the Situational Judgment Inventories. When fully operational, candidates' completed inventory booklets may be mailed to a central scoring location where responses are keypunched, automatically scored, and interpretive profiles printed by computer and returned to the site of testing. Turnaround time between date of testing and date of receiving completed results typically should be about 3 to 4 days at most.

*4. Regional assessment center materials.* Assessor's manuals have been prepared for each of the four police functions. The manuals contain complete definitions for the assessment dimensions taped by the simulations; detailed instructions for assessors on how to conduct exercises, observe candidates, and evaluate their performance; sample copies of all written instructions and materials for candidates; and sample copies of all report and rating forms to be completed by assessors after each exercise. The manuals were designed to be used by persons serving as assessors when police assessment centers become operational. For convenience, pages are color-coded according to whether they are explanations and instructions for assessors, instructions and other written materials to be used by candidates, or rating forms to be completed by assessors.

*5. Implementing regional assessment centers.* Centers may now be established on a regional basis to offer police assessment services periodically to local departments in the area. At first, a cadre of persons should be trained in each region who may be called upon periodically to serve as staff members. Our experience with the pretesting suggests that an effective approach is to include as staff members a mix of police officials, psychologists (who may be recruited from nearby schools, colleges, or universities), and local citizens active in some facet of social or community service work.

Persons who are to serve as assessors will require thorough training involving such things as:

- Learning the definitions of assessment dimensions to be measured during the assessment center.
- Reviewing the content and procedures of all exercises described in the Assessor's Manuals.
- Practicing role-play exercises and becoming familiar with the standardized role sets to be assumed by assessors.
- Practicing conducting background interviews.
- Practicing recording *behavioral* observations of candidates' performance and making evaluative ratings of performance backed by behavioral observations.
- Observing experienced assessors as they conduct an actual assessment center, carry out role-play exercises and background interviews, make behavioral observations of candidates, and rate their effectiveness.

Such training for assessors will probably require two or three days plus the time required to observe an actual assessment center being conducted by experienced assessors.

After an assessment center, staff members will typically pool their observations of candidates' performance and discuss candidates' overall effectiveness in the assessment dimensions. Assessors should discuss each other's observations and ratings for a candidate and reach consensus on how effective the candidate was on each dimension. The information may then be summarized in the form of a two- or three-page written report on each candidate's performance and his or her estimated potential for success in the particular police function being evaluated. The report will be sent directly to the candidate's local department where it may be combined with all other personnel information to derive a final judgment about him or her.

## D. Combined Operational Use of Police Career Index and Regional Assessment Centers for Personnel Decisions and Career Guidance of Police Personnel

We have described the development and validation of POLICE CAREER INDEX inventories and the development and pilot testing of assessment center job simulations for each of four police functions. These materials are now ready and available for operational use by local departments to help them in making selection, training, promotion, job transfer, and career counseling decisions for police candidates and/or experienced police officers already working in their departments.

Figure 1 shows how the POLICE CAREER INDEX inventories can be used in concert with Regional Police Officer Assessment Centers to provide information to local departments for use in carrying out the above personnel decision practices. The various stages shown in Figure 1 are described briefly below (The numbers on the boxes shown in Figure 1 refer to the stages discussed below.)

1. A local department seeking qualified candidates for vacancies would announce the availability of such positions and designate a date for administration of the POLICE CAREER INDEX (PCI) Biographical and Personal Information Inventory.
2. As candidates appear, the local department would carry out some form of preliminary screening, such as brief interviews, reference checks, etc., to de-

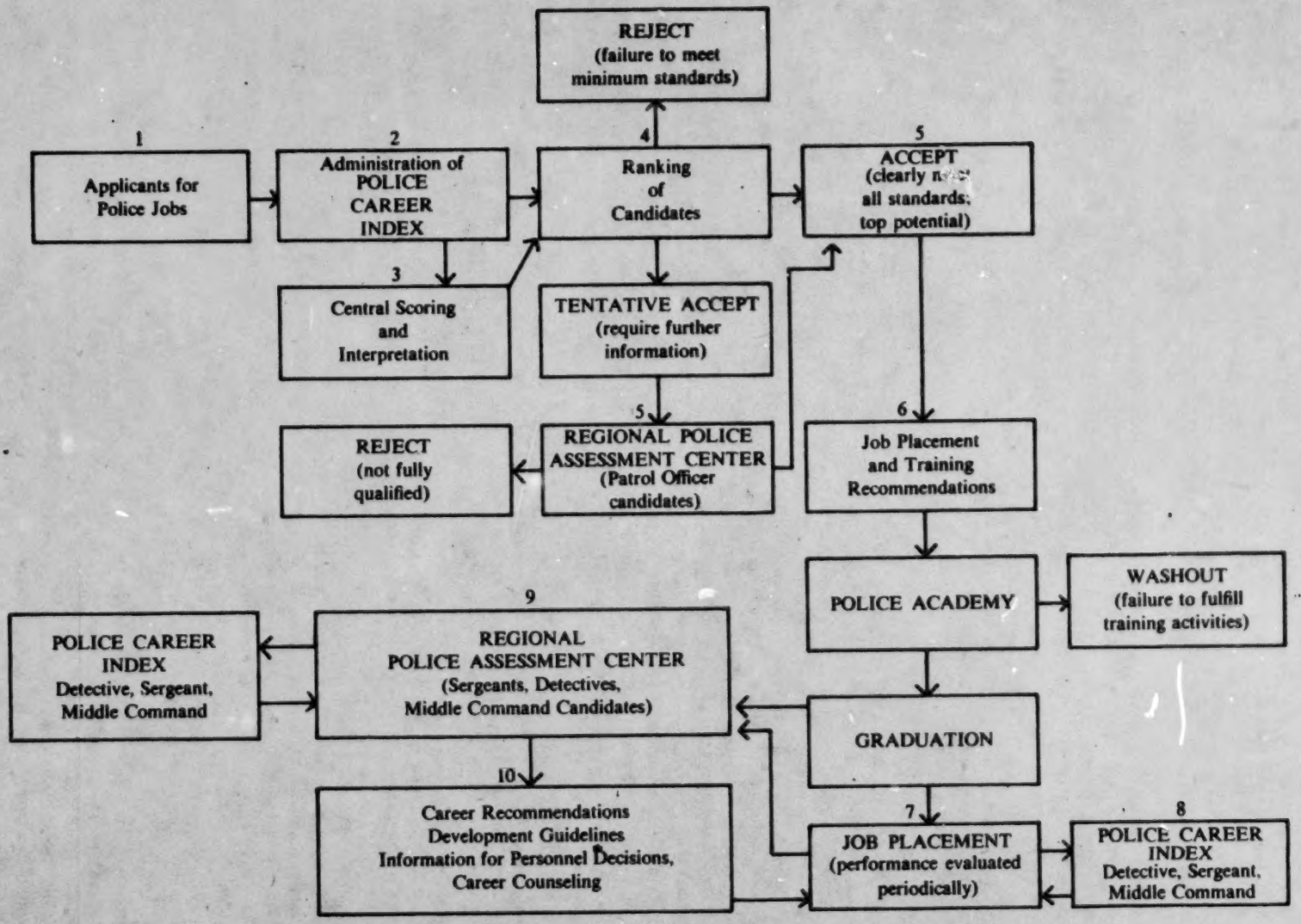



FIGURE 1. Stages of personnel selection, performance evaluation, assessment, and career placement decisions.

velop a roster of candidates to be admitted to the PCI inventory administration session. Sufficient numbers of PCI booklets and answer sheets would be obtained; and, on the appointed day, all qualified candidates would complete the PCI inventory.

3. Completed answer sheets would be forwarded to a centralized computer scoring service for scoring and automatic interpretation. At that time, the local department would also indicate the types of predictor scale scores desired for each candidate. Usually a department reviewing inexperienced candidates would probably request scoring for only the patrol officer predictor scales, although early guidance and training of a newly hired trainee could perhaps be enhanced by obtaining information relevant to the other police functions as well.

   Occasionally, a department might wish to evaluate an experienced police officer from some other department who might be under consideration for a supervisory or command job. At such times, the department would administer the appropriate PCI Situational Judgment Inventory (sergeant or middle command) in addition to the PCI Biographical and Personal Information Inventory and request scoring on the appropriate keys.

   The scoring service would carry out the scoring and return interpretative profiles (similar to those shown as Figures 3, 4, 5, and 6 in Chapter 2) to the local department within 2 days.

4. The PCI results would then be used in conjunction with additional departmental screening methods (such as physical examination, panel interview, etc.) to form an overall ranking of candidates.

5. The overall ranking of candidates would, for most departments, probably lead directly to the acceptance of the highest ranking persons to enter police training.

   Some departments might, however, seek further information on some of the candidates. These candidates would be asked to attend a one-day Regional Patrol Officer or Detective Assessment Center. This option would, of course, involve considerably greater cost to the department. But in some instances, the richness of the behavioral information might warrant obtaining such information.

6. Each candidate finally accepted should then be interviewed in a "feedback" session where the wealth of information obtained about her or him during the selection process would be discussed, with particular emphasis placed upon implications of the information for special training needs, areas of strength and weakness, possible career opportunities, etc.

7. Upon graduation from police academy training, officers would typically be assigned directly to police jobs either as patrol officers or as detectives. Again, some departments might seek further information about an officer early in his or her career by asking him or her to attend a regional police assessment center. More typically, however, the job performance of young officers would be evaluated periodically according to the department's existing personnel practices.

8. At some point in a young officer's developing career, the department might administer various inventories of the POLICE CAREER INDEX and request scoring on predictive scales bearing on detective, sergeant, and middle command jobs. Such information would then be used in conjunction with the accumulating knowledge of the officer's job performance to build a preliminary roster of "above average potential" officers for later promotional consideration when advancement opportunities develop.

9. At some stage (or, perhaps on several occasions) in an officer's career, the department would ask that he or she attend a regional police assessment center. There, the officer's potential would be evaluated according to the simulations and behavioral observation methods described in Chapter 3. If PCI information were not available on a candidate at the time of attending the regional center, the information would be obtained at that time as part of the total assessment center procedure.

10. Information about the officer's performance at the regional center would then be transmitted to appropriate persons in the department. It is desirable that the same information be given, with departmental approval, to the officer so that optimal use would be made of it in developing jointly agreed upon career recommendations, guidelines for further personal development, and basic behavioral information to be combined with all other information in developing a firm evaluation of his or her potential for serving effectively in the whole range of police positions available in the department.

The two "instruments" developed in this research program provide a total personnel evaluation system for decision making at all levels of police work, ranging from entry to the force to career guidance for individual officers and the development of increasingly accurate estimates of potential as those officers gain maturity and experience in their jobs. The central and most important

feature of these two coordinated assessment methods (the POLICE CAREER INDEX and the job-focused police officer assessment simulation exercises) is their grounding, at every stage of development, in behaviorally explicit patterns defined by the most critical features of police officer job performance. Thus, the POLICE CAREER INDEX scales are directly interpretable according to expected patterns of police job performance instead of being in the form of the usual traits measured by psychological tests. And, the assessment simulations have been shown to elicit successfully exactly those behavior patterns discovered to be critical to successful performance in different police jobs. The total evaluation system shown in Figure 1 provides a base, therefore, for gathering and accumulating information over time that is increasingly relevant and accurate for making personnel decisions about persons and jobs in police departments.

# CHAPTER I. BACKGROUND, RESEARCH NEEDS, AND RESEARCH APPROACH

## A. Background

Turbulent events over the last 15 years have forced upon the American public a profound awareness of the effectiveness or lack of effectiveness of its law enforcement agencies. In their recent book titled *Police Personnel Administration*, Stahl and Staufenberger (1974) comment that of all issues facing contemporary law enforcement, the one that is most central and of absolutely critical importance involves the quality, motivation, and utilization of the men and women who comprise local police agencies.

The principal objective of the extensive research program reported here has been the development and validation of a series of personnel assessment procedures to aid in evaluating applicants for police work and for developing programs of career guidance and upgrading for persons already in police work.

A first step in our research program was to learn as much as possible about the state of the art in police selection and promotion practices. A comprehensive review of the research literature (Groner, Johnson, & Dunnette, 1972, Heckman, Groner, Dunnette, & Johnson, 1972) yielded the following conclusions:

- First, much of the research has been rather spotty and piecemeal; no single "best" procedure for selecting police officers was identified. Some research studies (notably, Baehr et al., 1968; and, Furcon et al., 1971) have yielded potentially useful test validities, but the relative utilities of total selection systems have not been evaluated.
- Second, the focus of criterion research about the parameters of police effectiveness has been only weakly oriented toward actual dimensions of job performance. Most studies have relied upon criteria such as existing departmental ratings or global measures (such as rankings) of overall effectiveness. In many studies, therefore, raters have been left free to use their own potentially biased definitions of what constitutes police officer effectiveness, with the result that the usual sources of rating error (leniency, halo, etc.) probably have been quite prevalent.
- Finally, directions for next steps in police selection research were clearly evident. They included thorough job analyses and the use of critical incidents methodology to formulate behaviorally anchored rating scales (Landy & Goodin, 1974) for use in evaluating officers' levels of job effectiveness.

A more recent overview of current practices, strategies, and research needs in police officer selection (Eisenberg & Murray, 1974) also points up the critical importance of careful job analyses and better criterion development as a necessary basis for assessing future success in developing improved evaluation and personnel decision procedures for police applicants and police officers.

## B. Research Needs, Early Results, and Research Approach

Though our literature review was helpful, it needed to be supplemented with knowledge of how police personnel selection practices and problems may be viewed by practitioners. To this end, we interviewed top officials in 32 police departments in small, medium, and large cities throughout the country. (See Heckman, Groner, Dunnette, & Johnson, 1972, for a summary of interview findings.) Literature review and interview results formed the basis for a series of planning sessions directed toward formulating a broad research design for developing new police personnel evaluation and selection procedures. A central feature of the research plan came to be an emphasis on studying the critical features of four different police jobs: General Patrol Officers, Patrol Sergeants, Investigators or Detectives, and Intermediate Command Personnel (Lieutenants and Captains). Defining the critical dimensions of job performance for each job area would yield information that could be used in two distinct but importantly interrelated ways:

- First, critical incidents of actual police performance could form the basis for designing police job simulations or situational exercises which could be used to design assessment centers for evaluating the job potential of candidates.
- Second, such critical incidents of actual police performance could be used to define dimensions of job

performance effectiveness and to design behaviorally anchored rating scales for use in evaluating current job performance effectiveness of police officers. These rating scales could be used in a concurrent validation study to discover test and inventory item responses related to different levels of effectiveness for various performance dimensions.

In essence, then, our research plan was directed toward developing two types of "instruments" or procedures for use in evaluating candidates for each of four police jobs. One would be a brief (about two hours), easily administered and objectively scored inventory (the POLICE CAREER INDEX) to be used in the preliminary screening of applicants for police jobs. Our research plan was to examine inventory responses according to different patterns of rated performance on the behaviorally anchored scales and to combine valid items into behaviorally relevant scoring keys. Thus, scores on the POLICE CAREER INDEX would be interpretable directly according to expected patterns of job performance instead of being in the form of the usual traits measured by psychological tests and inventories. The second type of evaluation procedure would be in the form of a series of police job simulations to be used in police assessment centers. The police assessment centers would be available on a regional basis for use in evaluating the relative career potential, strengths and weaknesses, and training or developmental needs for police job candidates and/or currently employed police officers.

Figure 1 shows how the POLICE CAREER INDEX might be used in concert with regional police assessment centers to aid local departments in making selection, training, promotion, job transfer, and career counseling decisions for police candidates and/or experienced police officers.

The following steps were identified as being necessary for developing the two types of evaluation procedures shown in Figure 1:

1. Learn critical behaviors associated with effective and ineffective job performance in each of four police department jobs (Patrol Officer, Detective, Sergeant, Middle Command).
2. Develop behaviorally anchored job performance rating scales for each of four police department jobs.
3. Design and pretest performance samples (job simulations) to tap aspects of critical job performance dimensions for each of four police department jobs.
4. Validate and standardize inventories, tests, questionnaires, and personal history information for use in estimating likely patterns of job performance in each of four police department jobs (PCI).
5. Design and pilot test for operational readiness the specific procedures for conducting regional police assessment centers for each of four police department jobs.

The first two research steps were accomplished during the first year's research under Grant Award No. NI71–085–G. These activities and results are described in detail in the technical report by Heckman, Groner, Dunnette, and Johnson (1972). Step 4, the development of the PCI inventories, is described in Chapter 2 of the present report. Steps 3 and 5, the design and pretesting of simulations and the design and pilot testing of the assessment centers, are described in detail in Chapter 3 of the present report, Recommendations for operational use of the PCI and for implementing regional police assessment centers are given in Chapter 4 of the present report.

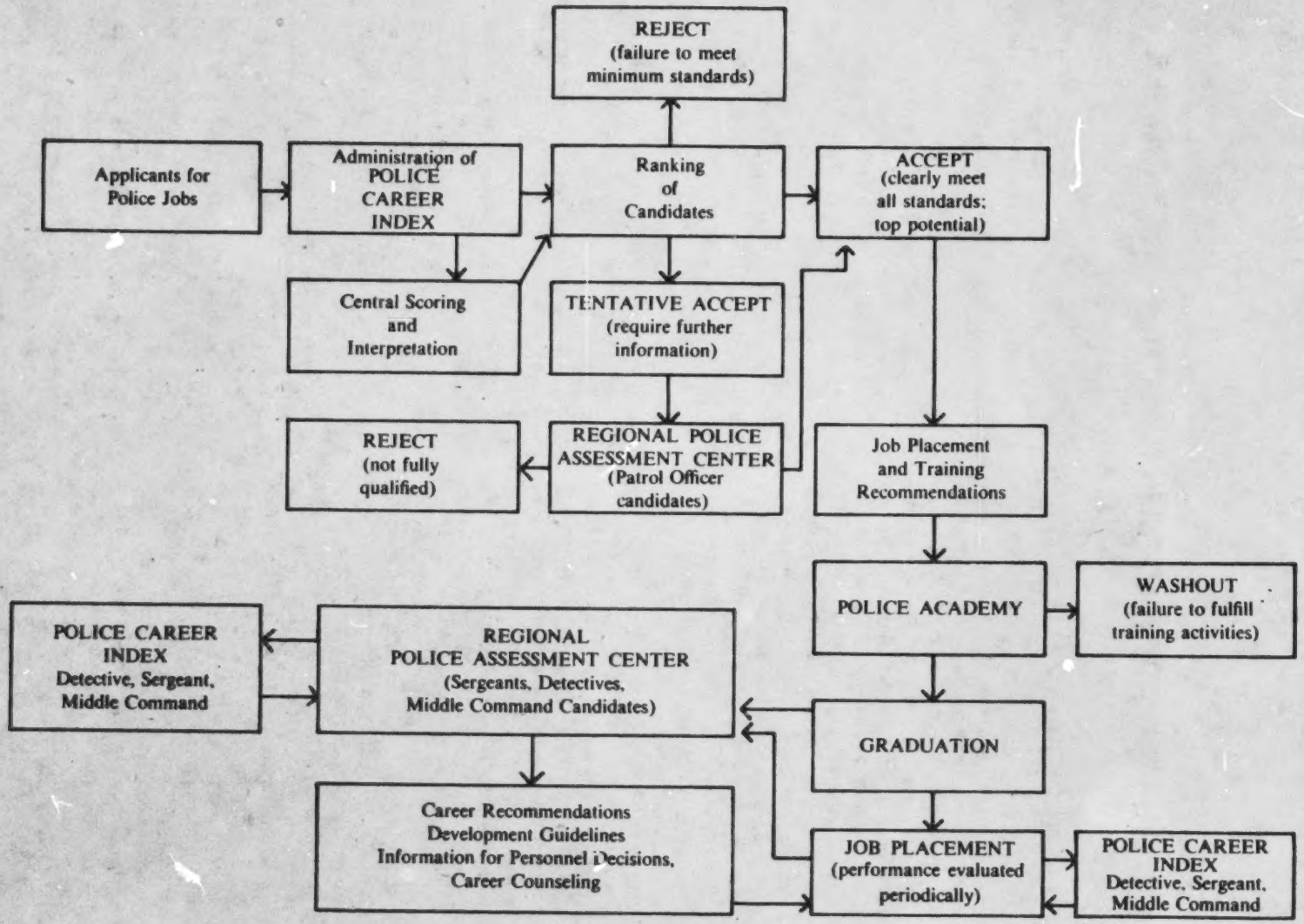



FIGURE 1. Stages of personnel selection, performance evaluation, assessment, and career placement decisions.

# CHAPTER II. DEVELOPMENT OF THE POLICE CAREER INDEX (PCI)

## A. Sequence of Development

As mentioned in Chapter 1, Step 4 of our research plan called for validating inventories, questionnaires, tests, and personal history information to develop four standardized, relatively brief, and easily administered inventories for use in the preliminary screening of candidates for each of four police jobs: Patrol Officer, Detective/Investigator, Sergeant, and Middle Level Command Personnel. We decided that the basic design of these four validation studies would be concurrent validation instead of predictive validation. (See Dunnette (1966) for a discussion of concurrent and predictive validation designs.)

The latter design would have required that applicants for each of the four police jobs be administered an experimental battery of tests and inventories, that some of the applicants be selected without regard to results on the tests, and that their actual job performance be evaluated after they had been on the job for some period of time. (Probably a year or more would be required before much confidence could be placed in job performance evaluations of these newly placed officers.) Only at the later time, after performance had been evaluated, could analytic procedures be undertaken to select the most valid items to comprise the final forms of the four PCIs. We did not want to suffer such a delay in our research program. Therefore, we determined to work with experienced officers. "Experienced" was defined to include officers who had been in a particular police job for over one year but less than five. By setting an upper limit of five years for each job area, we hoped to avoid drawing into our samples those officers who were much older than typical candidates for each of the four police jobs. We believed that inclusion of long-term officers or "old-timers" in our sample groups would introduce undesirable, age-related variance into both the job performance evaluations and the responses to items of the experimental tests and inventories. Unfortunately, we were not able in all instances to satisfy the intended constraints on length of service. Thus, officers with more than five years of experience did participate in our research in certain of the cooperating departments.

Having decided to use the basic design of concurrent validation, the sequence of steps necessary for developing the four PCIs became clearcut. These steps are listed below:

- Select or develop experimental tests and inventories to be administered to police officers.
- Secure cooperation from police departments so that broad geographical representation is assured. (Our initial goal was to obtain completed test and job performance information from 75 patrol officers, 750 detectives, 400 sergeants, and 400 middle command officers.)
- Administer experimental tests and inventories to police officers in the various cooperating police departments. Gather job performance descriptions (ratings) on behaviorally anchored scales from one or more senior officers for each of the officers tested with the experimental batteries.
- Analyze criterion (job performance ratings) information for officers in each job area according to interrater agreement and dimensionality of behavior descriptions. Choose the best criterion ratings to be used for analyzing test and inventory responses for each job area.
- Conduct item analyses separately for each inventory against each criterion dimension within each job area. Select most valid items and combine them to form "predictor" scales. Estimate cross-validities for each of the predictor scales.
- Examine predictor scale validities separately by police department and for persons subgrouped according to variables such as race, sex, age, etc.
- Develop standardized scoring and interpretive systems for reporting results for each of the four PCI inventories.

These seven steps are described in the following pages.

## B. Selection and Development of Experimental Tests and Inventories

1. *Critical performance dimensions.* The comprehensive study of critical performance dimensions for the four police jobs during our first year's research provided the

basic knowledge necessary for inferring the kinds of tests and inventories likely to be most predictive of police performance. Since each performance dimension was defined according to actual critical behaviors or incidents, our research staff could actually select and/or write items hypothesized to be related to specific behavioral incidents/dimensions/areas. The dimensions for each of the four jobs are listed below. Dimension titles, their general definitions, and the scales defining them are given in Appendix A of this report. The reader who wishes to review the hundreds of incidents that were actually gathered should refer to Heckman, Groner, Dunnette, and Johnson (1972).

*Patrol Officer dimensions*. The 11 dimensions used to describe the patrol officer's job included the following:

a. Crime Prevention
b. Using Force Appropriately
c. Traffic Maintenance and Control
d. Maintaining Public Safety
e. Investigating, Detecting, and Following Up on Criminal Activity
f. Report Writing
g. Integrity and Professional Ethics
h. Dealing Constructively with the Public
i. Handling Domestic Disputes
j. Commitment, Dedication, Conscientiousness
k. Team Work

*Detective/Investigator dimensions*. The 12 dimensions used to describe the detective/investigator job included the following:

a. Investigating the Scene of a Crime
b. Arrest, Search, and Seizure
c. Interrogating Suspects
d. Investigating a Case
e. Developing and Utilizing Informants
f. Report Writing and Paperwork
g. Appearing in Court
h. Public Relations
i. Dealing with Juveniles
j. Cooperating with Other Officers and Divisions
k. Conscientiousness and Dedication
l. Integrity and Honesty

*Sergeant dimensions*. The eight dimensions used to describe the sergeant's job included the following:

a. Concern for Subordinates
b. Scheduling, Coordination, Deployment, and Manpower Allocation
c. Supervision
d. Performing Administrative and Inspection Functions
e. Decision Making and Initiative Where No Firm Guidelines Exist
f. Training and Planning
g. Integrity, Dedication, and Conscientiousness
h. Dealing Effectively with the Public and Superiors

*Middle Command dimensions*. The nine dimensions used to describe the middle command job included the following:

a. Administrative Duties
b. Communications
c. Scheduling
d. Training
e. Supervision
f. Commending, Disciplining, and Assigning Efficiency Ratings
g. Field Command Situations
h. Public and Community Relations
i. Dedication, Integrity, Setting an Example

## C. Assignments for Test Item Selection and Writing

Seven research staff members were assigned the task of systematically examining the performance dimensions and writing items or selecting items from existing tests and inventories to tap those dimensions. Various types of tests or items also were specified as follows:

- Personal history or biographical items
- Personality items drawn from widely used inventories, such as the *Minnesota Multiphasic Personality Inventory, California Psychological Inventory,* etc.
- Likes and Dislikes or preference type items such as are contained in widely used vocational interest inventories. In addition to choosing items from existing inventories, special items were written that seemed to be particularly focused on preference areas inferred from the study of specific performance dimensions. An example of a series of ten preference items written to tap different aspects of a patrol officer's job is shown in the following:

- "Mark (X) which three of the following you LIKE BEST and which three you LIKE LEAST.

| LIKE BEST | LIKE LEAST | |
|---|---|---|
| _____ | _____ | giving orders and getting things moving |
| _____ | _____ | analyzing your own motives and feelings |
| _____ | _____ | working on a job that makes you feel that others depend on you whether or not you are on duty |

| | | |
|---|---|---|
| _____ | _____ | not having to work evenings and weekends |
| _____ | _____ | having a job where you don't have to get all fired up about it |
| _____ | _____ | having general directions on how to do a job but being able to work out the details yourself |
| _____ | _____ | thinking about how it would feel to be held at the point of a gun |
| _____ | _____ | reading detective stories |
| _____ | _____ | working with one other person |
| _____ | _____ | putting yourself in someone else's place and imagining how you would feel in the same situation." |

- Cognitive tests that had been shown to be useful in other studies such as Baehr et al. (1968), Furcon et al. (1971), or which might be inferred to be useful from a study of specific performance dimensions.
- Situational judgment items. The hundreds of actual critical incidents gathered as part of the development of performance dimensions were used to write items descriptive of situations an officer might encounter in carrying out his/her job. Alternative actions were developed, and the response format asked the respondent to choose the BEST and the WORST actions from the alternatives listed. An example of such a situational type item for a patrol officer's job is shown in the following:

- "While off duty, you observe a uniformed officer drinking in a bar. You would . . .

| | |
|---|---|
| ____BEST | 1. ignore the incident completely. |
| ____WORST | 2. find out who the officer is, whether or not he was on duty, and then call him and tell him to be more careful in the future about drinking while in uniform. |
| | 3. mention the incident to your shift supervisor and give him the man's badge number. |
| | 4. write down his badge number and report him to the chief's office." |

In developing these situational items, meetings were held with senior police officers and members of the Training Division of the Minneapolis Department to solicit their advice about the degree of realism shown by the items and the types of alternative actions that were neither obviously correct nor obviously wrong.

- Miscellaneous items. Various scales and inventories purporting to measure such areas as Internal-External Control, Dogmatism, Repressor-Sensitizor, Values, Self-Esteem, Needs, etc. were examined.

Staff members were assigned joint responsibilities for compiling various item types and for focusing on specific performance dimensions. Specific guidelines for the selection and writing of items were spelled out in a memo from Dunnette to staff members dated 25 September 1973. The memo read in part as follows:

"What constitutes 'good' items? They should be clear and easily understood—no big or unusual words, no double or triple negatives, no wordings that trap the respondent (such as, 'Have you stopped beating your wife?' or 'I wish I didn't worry so much about sexual matters.'). They should not tap identical information; that is, screen your items for overlap and redundancy. They should seem to be relevant to the behavior dimensions; or, if not, they should have proved useful in previous studies or be judged as potentially useful by some expert or be part of a scale that seems important. They should not be *obviously* related to race or sex, though this will, of course, be difficult to accomplish in many instances.

"Guidelines about numbers of items to be included are obviously difficult to give. But here's roughly what we might hope to have available for examination:

| | |
|---|---|
| Personal history— biographical ..... | 200–250 |
| Personality ..... | 400–500 |
| Likes and Dislikes ..... | 75–100 |
| Cognitive ..... | indeterminate |
| Situational Judgment .... | 100–125 [for each police job] |
| Miscellaneous ..... | 125–150" |

As soon as a preliminary battery of tests and inventories had been compiled for patrol officers, it was pretested with the help of 30 Minneapolis officers. Results of the pretest were as follows:

- Administration time was far too long. Nearly 6 hours were required instead of the intended 4 hours. This required dropping a number of the more esoteric scales from the inventories and the dropping of certain clerical type tests which had been initially introduced as a means of tapping spelling and grammatical knowledge.
- The initial situational judgment inventory was modified in accordance with comments made by the responding officers. The response format, which

had initially asked for a ranking of alternative actions, was also changed to the far simpler BEST-WORST format mentioned earlier. (This pretest experience, during which many of the officers complained about the items in the situational judgment inventory, was instrumental in causing us to solicit help from senior officers in the development of all other situational inventories.)

- Response distributions were examined for items of the Likes-and-Dislikes Inventory, the Biographical Inventory, and the Personality-type items. Items showing extreme response distributions (over 90 percent of respondents answering in the same way) were deleted from the inventories.

### D. Contents of Experimental Batteries

The above procedures yielded experimental batteries for administration to police officers. The tests and inventories comprising the batteries are described briefly as follows:

1. *Personal history/biographical inventory*. This inventory contained 149 items covering all aspects of background information. A sampling of areas covered and types of items includes:

- Personal characteristics, such as age, sex, race, height, weight, etc.
- Family background information, such as family size, relationships with siblings and parents, etc.
- High school accomplishments, experiences, and interests, such as scholastic achievement, courses liked or disliked, extracurricular activities, jobs held in high school, size of high school, etc.
- Descriptions of personal style, such as degree of aggressiveness, approach to solving problems, supervisory or leadership approach, degree of sociability, etc.
- Self-descriptive items, such as attitudes or opinions about hard work, risk taking, other persons' traits, desired goals, relative importance of various things to work and life satisfaction, etc.
- Several rather lengthy check lists of things that may or may not have happened to the respondent.*

An important feature of this biographical inventory was that we asked respondents to answer the questions according to their particular siutations at the time they were being considered for their present positions. The directions read as follows:

> "As you read each question, think back to the time when you first applied for your present position. Answer each question according to the way circumstances were then and not as they are now. In other words, assume that you are back to the time when you first applied for a position similar to the one you now have."

We hoped, of course, that this tactic would take us a step closer to discovering truly predictive relationships between personal history information and rated effectiveness in various police functions. Many aspects of biographical information change over the early years of a person's career (such as marital status, personal styles, attitudes, and opinions, etc.) Since the intended use of the PCI is for persons being considered for police jobs, it was important to obtain reports of biographical status prior to the effects of such career or time determined changes.

2. *Opinion and self-description inventories*. Personality type items were gathered from many existing personality inventories, such as the *Minnesota Multiphasic Personality Inventory*, the *California Psychological Inventory*, the *Guilford-Zimmerman* inventories, the *Jackson Personality Research Form*, etc. An initial selection of items was done by choosing those contained in scales shown in other studies to be correlated with police effectiveness. A secondary selection of items was accomplished by using our own clinical intuition to decide whether or not an item's content might be related to the various behavioral dimensions shown to be important in police jobs. As already mentioned, a small pretest among 30 Minneapolis police officers was used to delete items with extreme (90%-10%) response distribution splits. A total of 473 items was finally selected. Of these, 418 were in the form of a simple *True* or *False* response format. The remaining 55 requested a *True, Can't Say,* or *False* response. Two inventories were, therefore, pre-

*An example of one such check list is item 136, shown below:

136. Which of the following have you done? (Circle all that apply)

1. sold door to door (e.g., Fuller Brush man)
2. contract selling (encyclopedias, insurance)
3. worked in a gasoline station
4. clerked in a service establishment (such as a dry cleaner or shoe repair shop)
5. worked in a restaurant
6. sold in a retail store (such as a department store)
7. worked in a business office
8. conducted an interview
9. worked as a reporter
10. made a speech before a large group
11. none of the above

pared for the opinion and self-description items: Part I, a kind of warmup inventory containing the 55 three-response option items and Part II containing the remaining 418 items.

3. *Cognitive tests.* Only four cognitive tests were finally chosen for use in the experimental batteries. All are extremely short and brief (5 minutes) factor analytically derived measures of specific aptitudes. They are tests 1, 4, 7, and 10 of the Employee Aptitude Survey (Ruch & Ruch, 1963):

*Test 1: Verbal Comprehension,* a test of vocabulary knowledge; ability to use words in oral and written communication.

*Test 4: Visual Speed and Accuracy,* a test of speed and accuracy in perceiving details; requires the examinee to examine pairs of numbers and symbols and indicate which pairs are the same and which are different.

*Test 7: Verbal Reasoning,* a test of the ability to utilize facts in a logical way to derive a correct conclusion or to determine that no firm conclusion is possible from the facts presented.

*Test 10: Symbolic Reasoning,* a test involving cognitive processes similar to the verbal reasoning measure but utilizing symbols instead. Three sample items from Test 10 and their correct answers are shown below:

| | TRUE | FALSE | ? |
|---|---|---|---|
| 1. X = Y = Z; therefore, X = Z | ✓ | | |
| 2. X > Y > Z; therefore, X = Z | | ✓ | |
| 3. X ≠ Y ≠ Z; therefore, X ≯ Z | | | ✓ |

In addition to the above tests and inventories, different situational inventories (titled Situational Judgment Inventory) were developed for each of the four police jobs, and three different Likes and Dislikes inventories for patrol officers, detectives, and supervisory/management jobs (sergeants and middle command personnel). The situational judgment inventories contained 86 items for patrol officers, 90 for detectives, 97 for sergeants, and 88 for middle command personnel. The Inventories of Likes and Dislikes contained 205 items for patrol officers, 213 for detectives, and 195 for supervisory/management jobs.

## E. Administration of Tests and Collection of Performance Ratings

Police departments in nine cities located throughout the country agreed to cooperate with us in our collection of test and job performance information. A supplementary grant from the Law Enforcement Assistance Administration allowed us to compensate officers for time required for the test administration and for the time involved in completing performance ratings for subordinate officers. Most of the data collection was, therefore, carried out during off-duty hours for the participating officers. The data collection occurred over almost exactly a year's time extending from mid-October 1973 through mid-October 1974. During that time, some form of information (test and/or rated performance) was collected from or about over 2,000 police officers from departments in the cities of Albuquerque, Cincinnati, Des Moines, Miami (Dade County), Minneapolis, Portland (Oregon), San Diego, Tucson, and Washington, D.C.

Table 1 shows types of data collected, kinds of police officers represented, and dates for each of the nine cities. Table 2 shows the total numbers of police officers for whom information was finally obtained in each of the cities. Note that performance ratings were obtained for larger numbers of patrol officers than were tested in the two cities of Minneapolis and Portland. Our intention was to compare mean job performance ratings of the officers who took tests with the mean job performance ratings of all officers in the department. However, logistic difficulties prevented us from obtaining ratings for the entire population of patrol officers in either Portland or Minneapolis. Nonetheless, comparison of performance rating means for all tested and all rated officers provides a crude estimate of the degree of representativeness of the officers who were tested in these two cities. Comparisons are shown later in the section devoted to our discussion of criterion analyses. [Note that the numbers rated and the numbers actually tested differ slightly in several other cities as well. These discrepancies occurred simply because testing session records were not always entirely complete at the time the rating sessions were held. Therefore, the rosters for the rating sessions contained some names of persons who were not tested and, unfortunately, failed to contain a few who actually were tested.] Information in Table 2 shows that we reached our sample size goal only for patrol officers. We came close to the goal of 400 for sergeants, but we severely undershot the goals of 750 for detectives and 400 for middle command officers.

1. *Test administration.*

a. *General approach.* Before administering any tests, our test administrators (staff members from Personnel Decisions, Inc.) explained the overall purpose of the research study as thoroughly as possible. Special emphasis was given to the confidentiality of all results, to the nationwide character of the study, and to the importance of the part they were playing in helping to develop good guidelines for selecting future police officers. They

TABLE 1

*Data Collection Sequence According to Time, City, and Type of Police Jobs*

| City | Testing Dates | Functions Tested | Rating Dates | Functions Rated |
|---|---|---|---|---|
| Albuquerque | July 31–August 3, 1974 | D S M | October 10, 1974 | D S M |
| Cincinnati | February 3–8, 1974 | P D | February 3–8, 1974 | P D |
| | March 25–29, 1974 | S M | March 25–29, 1974 | S M |
| Des Moines | February 17–22, 1974 | P D | February 17–22, 1974 | P D |
| | March 10–15, 1974 | S M | March 10–15, 1974 | S M |
| Miami (Dade County) | March 2–6, 1974 | all | March 23–26, 1974 | all |
| | | | April 29–May 4, 1974 | all |
| | | | October 13–15, 1974 | all |
| Minneapolis | October 15–19, 1973 | P | October 23–24, 1973 | P |
| | February 12, 1974 | D | October 29, 1973 | P |
| | March 20–21, 1974 | S | April, 1974 | D |
| | April 9–10, 1974 | M | April 25, 1974 | S |
| | | | May 1, 1974 | M |
| Portland | November 25–30, 1973 | P | November 25–30, 1973 | P |
| | April 28–May 3, 1974 | D S M | April 28–May 3, 1974 | all |
| San Diego | August 4, 1974 | D S M | August 4, 1974 | D S M |
| Tucson | (Internal Testing) | D S M | October 9, 1974 | D S M |
| Washington, D. C. | April 7–11, 1974 | P D | May 19–23, 1974 | P D |
| | May 19–23, 1974 | S M | September 10–18, 1974 | all |
| | | | October 1–4, 1974 | all |

*Key:* P = Patrolmen; D = Detectives; S = Sergeants; M = Middle Level Commanders.

TABLE 2

*Numbers of Police Officers Tested* and Rated According to City and Type of Police Job*

| | Patrol Officers | | Detectives | | Sergeants | | Middle Command | | TOTALS | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Tested | Rated | Tested | Rated | Tested | Rated | Tested | Rated | Tested | Rated |
| Albuquerque | .. | .. | 8 | 8 | 9 | 9 | 1 | 1 | 18 | 18 |
| Cincinnati | 135 | 146 | 63 | 67 | 68 | 74 | 28 | 28 | 294 | 315 |
| Des Moines | 109 | 116 | 29 | 33 | 43 | 46 | 20 | 23 | 201 | 218 |
| Miami (Dade County) | 82 | 98 | 71 | 94 | 55 | 54 | 31 | 31 | 239 | 277 |
| Minneapolis | 190 | 238 | 48 | 54 | 40 | 40 | 31 | 28 | 309 | 360 |
| Portland | 81 | 141 | 52 | 62 | 16 | 18 | 20 | 23 | 169 | 244 |
| San Diego | .. | .. | 56 | 56 | 53 | 53 | 31 | 31 | 140 | 140 |
| Tucson | .. | .. | 22 | 22 | 3 | 3 | 3 | 3 | 28 | 28 |
| Washington, D. C. | 171 | 173 | 103 | 89 | 95 | 93 | 50 | 49 | 419 | 404 |
| TOTALS | 768 | 912 | 452 | 485 | 382 | 390 | 215 | 217 | 1817 | 2004 |

**Note:* In no city did exactly the same number of officers complete *all* tests of the battery. Because of time limitations, from five to ten fewer officers in some cities completed the Likes and Dislikes Inventory and/or the Opinion and Self-Description Inventory. The numbers above refer to the numbers of individuals tested, regardless of whether or not they completed *all* tests.

were told that no one in their department would ever have access to any results and that their test performance could in no way affect their current or future job status. Nonetheless, they were urged to do their best and to take the tests seriously in order to give an accurate portrayal of how an applicant might approach such a testing session. Finally, they were informed that they should not place their names anywhere on the tests. We explained that leaving their names off the tests would be a safeguard against the remote possibility of any tests being lost or misplaced. The tests and inventories of each battery were, however, numbered and each police officer wrote his name and test battery number on a roster. We explained that the numbers were necessary for three reasons: (1) to keep each person's tests and inventories together as a package; 2) to develop a roster of names for use in writing the checks to reimburse them for their time; and, (3) to allow us to develop further information about their job performance so that their testing results could be correlated with factors such as their amount of experience and overall "style" in their jobs. We did not explicitly mention the necessity of obtaining job perfor-

mance ratings for each participant. If anyone asked whether or not such ratings would be obtained, the administrator said that they would be and took the opportunity to reemphasize the fact of tight security and complete confidentiality of all such research information.

b. *Actual administration.* A specified order was used for administering the tests and inventories. The cognitive tests were administered first because they are the only ones with specific time limits. The tests of Visual Speed (EAS Test 4) and Verbal Comprehension (EAS Test 1) were administered first, because their formats and content are rather common and probably quite familiar to most persons. Beginning with relatively familiar tests seemed to be a good way of gaining early rapport with the examinees. The test of Verbal Reasoning (EAS Test 7) was administered next, followed by the test of Symbolic Reasoning (EAS Test 10).

The remaining inventories, stacked in a particular order, were then distributed to the police officers. The test administrators went over the directions for each of the inventories. After allowing time for questions and further clarification, the examinees were instructed to work at their own speed to complete all the inventories. In all cities but Washington, the testing was carried out in a single session lasting 4 or 5 hours. In Washington, the officers took the tests in two sessions of about 2½ hours each scheduled on two successive days. Each officer was paid $24 for participating in the test administration.

c. *Order of inventories.* The untimed inventories were arranged so that the officers would take them in the following order:

- *Biographical History Information.* This was placed first because it is quite easy and straightforward to fill in. The questions bearing on personal history are generally innocuous. We also expected that biographical information would provide a rich yield of valid items. Thus, it seemed best to have the officers complete that inventory while they were still "fresh," before test-taking fatigue had set in.
- *Situational Judgment Inventory.* Since the Situational Judgment Inventory contained items derived directly from actual job experiences reported by police officers, we believed its administration would help to maintain continued interest and motivation in the testing situation. The examinees could see the relevance of these situational items and would be less likely to be "turned off" by the testing session.
- *Opinion and Self-Description Inventory, Part II.* Officers were urged to take a coffee break after finishing the Situational Judgment Inventory. We placed the OSDI (Part II) in position to be taken immediately after the opportunity to relax over coffee. The OSDI was the longest inventory and, of course, contained potentially the most onerous and seemingly irrelevant content. Our administrators were entirely straightforward in warning the officers that this would be the most "burdensome" of the inventories, but that they should just "bear with it" and try not to dwell at length on any single response.
- *Inventory of Likes and Dislikes.*
- *Opinion and Self-Description Inventory, Part I.* These were placed last because they are shorter and relatively much easier to complete than the OSDI Part II. The preference inventory, in particular, is once again quite innocuous and easy to fill in. Thus, these two reasonably short, innocuous, and easy inventories were placed in position where the accumulated test-taking fatigue among the officers would have least effect on their attentiveness and overall rapport. [Though the OSDI, Part I, was initially designed to be a "warmup" inventory in preparation for the later, more lengthy tests, we found its use as a "wrapup" inventory to have greater utility.]

2. *Performance Ratings.*

a. *Rating materials.* The Job Performance Description Booklets containing rating guidelines, dimension definitions, and performance scales for each of the four police jobs are contained in Appendix A. As can be seen, the first four pages include: (1) a rather detailed statement about some of the difficulties inherent in rating job performance; (2) a listing of the particular performance dimensions to be used for the officers being rated; (3) instructions for recording ratings on a special work sheet; and, (4) a section detailing "things to guard against" or sources of error that can contribute inaccuracy to ratings.*

As implied above, in addition to the Job Performance Description Booklet, each rater received a set of rating work sheets, one for each of the performance dimensions to be rated. Figure 2 shows one such work sheet (for Sergeants' Job Category A: Concern for Subordinates). The use of this work sheet is explained in the following section.

b. *Obtaining ratings.* Rating sessions were scheduled with small groups of senior officers who had been designated by department officials as having good knowledge of the job performance of the persons to be rated. When ratings were obtained for supervisory personnel, sergeants, and especially middle command offic-

*The "things to guard against" section included comments designed to help the raters guard against halo error, making ratings based on limited observations, stereotype error, and tendencies to ignore the definitions provided for each of the job performance dimensions.

**CONCERN FOR SUBORDINATES**

Showing concern for subordinates; being considerate; taking personal interest in their problems; giving recognition when deserved; using human relations principles; filling in for subordinates who have special problems; being aware of subordinates' needs and feelings.

| OFFICERS RATED | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |

FIGURE 2. *Rating Worksheet for Sergeants' Job Category A: CONCERN FOR SUBORDINATES.*

ers, our staff member often conducted a rating session with just one or two persons—frequently in the senior officer's own office setting. By providing such individual attention and guidance ("hand holding"), we hoped to assure excellent rapport and attentiveness in the rating situation. Before handing out the rating materials, our staff member** gave a brief 15 or 20 minute talk about the project, its importance, its magnitude, and the absolutely crucial importance of accurate performance ratings to the overall success of the venture. In addition, the talk included comments about problems and sources of error in performance rating and a careful explanation of how we had developed special behaviorally anchored performance description scales in order to help overcome most of these sources of error.

The Job Performance Description Booklets were then distributed. The raters were asked to read the explanatory pages of their Job Performance Description Booklets and to study briefly the actual scales to be used in rating their subordinate officers. After allowing time for questions and further clarification, the rating work sheets were distributed. In some departments, we had been able to identify with some precision the subordinate officers who were to be rated by each senior officer. In those instances, we wrote the names of the persons to be rated in the appropriate blanks on each officer's set of rating work sheets. Rating officers were asked to read the names of the persons listed and to cross off the names of any whom they believed they did not know well enough to rate. In addition, they were asked to add the names of any subordinate whom they *did* know well and whom they wanted to rate. [We believed it wise to have such persons added even though they might not have been tested. In this way, each rating officer would be reporting on the entire range of job performance levels of all persons known to him rather than the potentially somewhat restricted range of just the persons who had taken the test batteries.]

It was not possible in many departments to identify ahead of time exactly which subordinate officers were to be rated by each senior officer. In those instances, rosters or persons who had taken the tests were distributed along with the rating work sheets. Rating officers were asked to examine the rosters and to write in the names of all persons whom they believed they knew well enough to rate. We were intent, of course, upon obtaining more than one rating for as many of the tested officers as possible in order to examine the degree of interrater agreement (reliabilities) on the various scales. Our practice of distributing rosters of examinees to the raters

**In some instances, PDI staff members trained department officials to conduct the rating sessions. In every case, they used an approach identical to the one described above.

allowed us to obtain more such multiple ratings than would otherwise have been the case. Rating officers were then instructed to consider the job performance dimensions one at a time, to rate all officers on a single dimension, to go on to the next dimension, rate all officers on it, and so on until all dimensions had been considered and all officers rated. They recorded their ratings simply by marking the appropriate blank on the nine-step scale for each of the officers being rated.

c. *Reactions to the rating process.* The rating process obviously was time-consuming and difficult. Nonetheless, the job performance scales were greeted almost uniformly with enthusiasm. Our strategy of defining different levels of performance with real incidents of police job behavior made this rating task more interesting and considerably more acceptable to the senior officers than is typically the case with seemingly simpler but less carefully defined rating scales.

As shown in Table 2, performance ratings were obtained for a total of 2,004 officers. As mentioned, however, ratings were obtained for many officers who were not tested. Similarly, they were *not* obtained for several officers who were tested. Thus, the actual number of officers for whom test responses could be compared with performance ratings was considerably below 2,004. These matters are dealt with in the sections that follow.

## F. Analysis of Performance Ratings

1. *Overview of analysis and findings.* In spite of the pains taken to obtain the best possible information about police officers' job performance, statistical analyses were required to examine the overall quality of the ratings actually obtained. Two major sets of analyses were carried out.

First, interrater agreement was examined for each rating dimension separately according to ratings obtained from each of the cooperating departments. As is detailed below, these analyses of the reliabilities of the ratings showed large differences from city to city and between various performance dimensions in the relative quality of the ratings. As a result, decisions were made (described in detail below) to delete certain performance ratings from further analysis.

After discarding the ratings of low quality, the second major form of statistical analysis was carried out. This consisted of an examination via factor analysis of the dimensionality of the performance rating correlation matrices for the four police jobs. Had we been entirely successful in developing conceptually distinct job performance dimensions, we could expect relatively low intercorrelations between the various dimensions on which the officers were rated. On the other hand, to the extent that the dimensions were either not distinct or the defining scales incapable of yielding ratings free from halo error, a dimensional analysis would show only a single general factor. As is detailed below, these factor analyses showed that halo error was *not* overcome with any satisfying degree of success. Only a few of the job performance dimensions for each of the police jobs appeared to be sufficiently distinct statistically to justify their retention for further analysis.

Methods of analysis, results obtained, and decisions made concerning the criterion data are described in the two following sections: Reliability Studies and Dimensional Analyses.

2. *Reliability studies.* Our study of the interrater agreement for various scales and different departments used ratings for all persons as opposed to using only those for persons with test information. We reasoned that estimates based on the broader range of persons would be more representative of the scales in actual use than those based on ratings for only those persons who had been tested.

Interrater reliability coefficients were computed by estimating the ratio of differences between raters rating the same subordinates to differences for raters rating different subordinates. Essentially, this involves computing the ratio between rater agreement and rater disagreement and is a variant of the intraclass correlation coefficient (Ebel, 1951). In a general way, the measure of interrater agreement can be expressed as:

$$1 - \frac{\text{Interrater variance within ratings given to jointly rated ratees}}{\textit{Variance of mean ratings given to all ratees}}$$

Computationally, the numerator of the ratio in the above expression is the pooled within-set variance across all groupings of jointly rated subordinates. The denominator is simply the total variance across all ratees' mean ratings or dimension "scores"—whether based on just one rating or on multiple ratings. If all raters agree completely for every jointly rated subordinate, the value of the ratio becomes 0 and the reliability will equal 1. If the differences between different raters' evaluations of the same person are as large as the differences between the ratings given to all persons, the index will equal 0, the scale is completely unreliable, and our confidence in the statistical stability of any rated differences between two or more persons is put to a severe test. If rated performance differences between subordinates are very large relative to the average amount of disagreement for an individual subordinate, the errors are sufficiently small so that they are unlikely to affect the general stability of the ordering of the ratees, and the ratings can be used with confidence to distinguish effective performers from less effective performers for test validation

purposes. The above line of reasoning outlines our rationale for examining the interrater reliabilities of the performance ratings as a way of identifying and discarding from further analysis any ratings of such low quality as to cause us to seriously question their usefulness for test validation purposes.

a. *Patrol officer ratings.* Table 3 shows the numbers of patrol officers who were rated by various numbers of raters. Needless to say, we were surprised to note that a few officers had been rated by so many persons. It seemed unlikely that as many persons as five or six would be equally well qualified to rate the job performance of any single patrol officer. We suspected that the reliabilities of ratings for persons with high numbers of multiple ratings would, therefore, be low. The reliabilities of multiple ratings were examined separately according to city, and our expectations generally were confirmed. For example:

- Reliabilities for nine Washington officers with five or more ratings were uniformly low.
- Reliabilities for four Cincinnati officers with five or more ratings were uniformly low.
- Reliabilities for 11 Portland officers with three ratings were uniformly low.
- Reliabilities for two Des Moines officers with three ratings were uniformly low.

Because we have no good way of identifying which raters were the more qualified to rate the above 26 officers, we decided simply to delete their performance ratings from any further analyses.

Tables 4–10 show estimates of reliabilities for each of the patrol officer dimensions for all cities together and separately for each of the cities. As will be noted, three estimates are given for each dimension. The estimates in the first column of each of the tables are computed from all available information, including officers with just one rating, as well as those with multiple ratings. This means that the pooled within variance across jointly rated subordinates (numerator) is based on fewer subjects than the between variance based on all ratees' mean "scores" (denominator). The reliability estimates in the first column are the best estimate for the stability of the actual "scores" for each officer on each dimension, but they

TABLE 3

*Numbers of Patrol Officers Whose Job Performance Was Rated by 1, 2, 3, Etc. Raters*

| Number of Raters | Number of Ratees |
|---|---|
| 1 | 329 |
| 2 | 288 |
| 3 | 132 |
| 4 | 104 |
| 5 | 45 |
| 6 | 7 |
| 7 | 2 |
| 8 | 3 |
| 9 | 1 |
| 10 | 0 |
| 11 | 1 |
| | 912 |

TABLE 4

*Interrater Reliability Estimates* for Job Performance Ratings for All Patrol Officers (N = 886**)*

| Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=288) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|
| A Crime Prevention | 53 | 53 | 50 |
| B Use of Force | 48 | 44 | 46 |
| C Traffic Control | 47 | 49 | 45 |
| D Public Safety | 43 | 39 | 41 |
| E Investigation | 48 | 47 | 46 |
| F Report Writing | 50 | 53 | 48 |
| G Ethical Conduct | 37 | 29 | 34 |
| H Dealing with Public | 37 | 48 | 35 |
| I Domestic Disputes | 37 | 26 | 35 |
| J Commitment | 55 | 55 | 53 |
| K Teamwork | 53 | 48 | 51 |
| L Overall Rating | 67 | 72 | 65 |
| Sum of All Ratings | 64 | 66 | 61 |
| Mean Number of Raters (K) | 2.20 | 2.00 | .. |

*Decimals omitted.

**Twenty-six officers with extremely unreliable multiple ratings have been deleted from this analysis.

TABLE 5

*Interrater Reliability Estimates* for Job Performance Ratings for Cincinnati Patrol Officers*
*(N = 142**)*

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=74) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Crime Prevention | 52 | 66 | 49 |
| B | Use of Force | 36 | 29 | 34 |
| C | Traffic Control | 55 | 58 | 52 |
| D | Public Safety | 31 | 31 | 28 |
| E | Investigation | 52 | 66 | 49 |
| F | Report Writing | 50 | 57 | 47 |
| G | Ethical Conduct | 24 | −12 | 22 |
| H | Dealing with Public | 45 | 56 | 42 |
| I | Domestic Disputes | 39 | 30 | 37 |
| J | Commitment | 61 | 68 | 59 |
| K | Teamwork | 53 | 58 | 51 |
| L | Overall Rating | 63 | 81 | 60 |
| | Sum of All Ratings | 62 | 72 | 59 |
| | Mean Number of Raters (K) | 2.25 | 2.00 | .. |

*Decimals omitted.

**Four officers with extremely unreliable multiple ratings have been deleted from this analysis.

TABLE 6

*Interrater Reliability Estimates* for Job Performance Ratings for Des Moines Patrol Officers*
*(N = 114**)*

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=51) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Crime Prevention | 23 | 33 | 29 |
| B | Use of Force | 29 | 52 | 36 |
| C | Traffic Control | −38 | −03 | 00 |
| D | Public Safety | 13 | 35 | 18 |
| E | Investigation | −42 | −16 | 00 |
| F | Report Writing | 08 | 29 | 11 |
| G | Ethical Conduct | −31 | −03 | 00 |
| H | Dealing with Public | 20 | 51 | 25 |
| I | Domestic Disputes | −23 | −24 | 00 |
| J | Commitment | −33 | −16 | 00 |
| K | Teamwork | −06 | 02 | 00 |
| L | Overall Rating | 27 | 37 | 34 |
| | Sum of All Ratings | 01 | 25 | 01 |
| | Mean Number of Raters (K) | 1.45 | 2.00 | .. |

*Decimals omitted.

**Two officers with extremely unreliable multiple ratings have been deleted from this analysis.

TABLE 7

*Interrater Reliability Estimates* for Job Performance Ratings for Miami (Dade County) Patrol Officers (N = 98)*

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=48) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Crime Prevention | −41 | −27 | 00 |
| B | Use of Force | −05 | 14 | 00 |
| C | Traffic Control | 31 | 39 | 37 |
| D | Public Safety | 14 | 13 | 17 |
| E | Investigation | 03 | 15 | 04 |
| F | Report Writing | 26 | 37 | 32 |
| G | Ethical Conduct | 42 | 37 | 49 |
| H | Dealing with Public | 51 | 54 | 58 |
| I | Domestic Disputes | 10 | −19 | 13 |
| J | Commitment | 28 | 14 | 34 |
| K | Teamwork | 09 | −04 | 12 |
| L | Overall Rating | 73 | 76 | 78 |
| | Sum of All Ratings | 48 | 48 | 56 |
| | Mean Number of Raters (K) | 1.50 | 2.00 | .. |

*Decimals omitted.

TABLE 8

*Interrater Reliability Estimates* for Job Performance Ratings for Minneapolis Patrol Officers (N = 238)*

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=29) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Crime Prevention | 64 | 30 | 53 |
| B | Use of Force | 68 | 84 | 57 |
| C | Traffic Control | 48 | −07 | 36 |
| D | Public Safety | 45 | 29 | 34 |
| E | Investigation | 56 | 64 | 44 |
| F | Report Writing | 53 | 64 | 41 |
| G | Ethical Conduct | 47 | 25 | 35 |
| H | Dealing with Public | 44 | 11 | 33 |
| I | Domestic Disputes | 51 | 27 | 39 |
| J | Commitment | 56 | 73 | 44 |
| K | Teamwork | 49 | 56 | 37 |
| L | Overall Rating | 70 | 78 | 59 |
| | Sum of All Ratings | 70 | 70 | 59 |
| | Mean Number of Raters (K) | 3.25 | 2.00 | .. |

*Decimals omitted.

TABLE 9

*Interrater Reliability Estimates* for Job Performance Ratings for Portland Patrol Officers (N = 130**)*

| Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=37) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|
| A Crime Prevention | 64 | 82 | 64 |
| B Use of Force | 47 | 35 | 47 |
| C Traffic Control | 62 | 61 | 61 |
| D Public Safety | 42 | 56 | 41 |
| E Investigation | 57 | 67 | 56 |
| F Report Writing | 46 | 60 | 46 |
| G Ethical Conduct | 25 | 46 | 25 |
| H Dealing with Public | 36 | 47 | 35 |
| I Domestic Disputes | 24 | 54 | 25 |
| J Commitment | 59 | 88 | 59 |
| K Teamwork | 51 | 66 | 50 |
| L Overall Rating | 70 | 84 | 70 |
| Sum of All Ratings | 67 | 78 | 66 |
| Mean Number of Raters (K) | 2.00 | 2.00 | .. |

*Decimals omitted.

**Eleven officers with extremely unreliable multiple ratings have been deleted from this analysis.

TABLE 10

*Interrater Reliability Estimates* for Job Performance Ratings for Washington Patrol Officers (N = 164**)*

| Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=49) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|
| A Crime Prevention | 31 | 62 | 33 |
| B Use of Force | 22 | 29 | 24 |
| C Traffic Control | 18 | 47 | 20 |
| D Public Safety | 14 | 21 | 15 |
| E Investigation | 13 | 23 | 14 |
| F Report Writing | 29 | 60 | 31 |
| G Ethical Conduct | 25 | 29 | 26 |
| H Dealing with Public | −07 | 19 | 00 |
| I Domestic Disputes | 00 | 00 | 00 |
| J Commitment | 42 | 43 | 44 |
| K Teamwork | 26 | 43 | 28 |
| L Overall Rating | 43 | 55 | 45 |
| Sum of All Ratings | 33 | 56 | 35 |
| Mean Number of Raters (K) | .. | 2.00 | .. |

*Decimals omitted.

**Nine officers with extremely unreliable ratings have been deleted from this analysis.

may be somewhat imprecise (probably somewhat overestimated) because of the computational anomaly mentioned above.

The estimates in the second column are computed for officers who were actually rated by exactly two raters. The estimates in the third column were derived by using the Spearman-Brown formula to correct the first column estimates to a base of K = 2; that is, a statistically corrected estimate based on the hypothetical condition where each ratee is assumed to have been rated by exactly two raters. City-to-city comparisons of estimated reliabilities for various dimensions are best made by comparing the estimates given in columns two and three, since they both have essentially the same base of two raters per ratee.

As can be readily noted from Tables 4–10, reliabilities are generally acceptable for all dimensions for the cities of Cincinnati, Minneapolis, and Portland. They are considerably poorer and many are clearly unacceptable for the departments in Des Moines, Miami, and Washington. We decided to discard from further analyses the ratings of lowest quality and to recompute overall reliability estimates across all cities for those sets of data that were most reliable. The following decisions were made:

- Ratings for Des Moines officers on dimensions C, E, G, I, J, and K were discarded.
- Ratings for Miami officers on dimensions A, B, D, E, I, J, and K were discarded.
- Ratings for Washington officers on dimensions D, E, H, and I were discarded.

New reliability estimates were then recomputed across all cities for the remaining information. In this instance, however, we computed reliabilities for only those officers for whom test information was available. We reasoned simply that reliability estimates should be obtained, at this stage, for only those ratings to be used as criterion scores in generating item validity statistics. Table 11 shows the final reliability estimates for tested patrol officers for each of the dimensions (with data discarded as specified above) across all cities. In general, the reliabilities are acceptable, though in some instances (notably for dimensions D, E, and I) quite severe reductions in sample size were necessary in order to achieve higher and more acceptable values.

b. *Detective ratings*. Multiple ratings were used much less frequently for detectives than for patrol officers. Table 12 shows the distributions of multiple ratings for detectives. Data for the two detectives rated by four persons were deleted from further analyses.

Reliability estimates were computed for each of the detective dimensions for all cities together and separately for each of the cities. In the interest of saving space, tables showing the separate reliability estimates for each city are given in Appendix B. However, Table 13 shows the reliability estimates for detective dimension ratings for all detectives taken together. Dimensions A, B, C, D, the Overall Rating, and the Sum of Ratings show reasonably good reliabilities. Fortunately, these dimensions are exactly the ones which seem intuitively to be most important in the overall job of investigation. In contrast, dimensions E, F, G, H, I, J, K, and L show low reliabilities. Examining the separate reliability esti-

TABLE 11

*Interrater Reliability Estimates* for Retained Job Performance Ratings for All Tested Patrol Officers*

| Dimension | Overall Estimate | | Estimate Based on Ratees with Two Ratings Only | |
|---|---|---|---|---|
| | N | r | N | r |
| A Crime Prevention | 618 | 57 | 202 | 58 |
| B Use of Force | 626 | 54 | 202 | 46 |
| C Traffic Control | 591 | 53 | 195 | 55 |
| D Public Safety | 465 | 46 | 153 | 41 |
| E Investigation | 363 | 60 | 106 | 67 |
| F Report Writing | 702 | 52 | 242 | 52 |
| G Ethical Conduct | 593 | 40 | 194 | 31 |
| H Dealing with Public | 542 | 48 | 193 | 59 |
| I Domestic Disputes | 360 | 53 | 106 | 28 |
| J Commitment | 522 | 64 | 154 | 66 |
| K Teamwork | 520 | 59 | 155 | 55 |
| L Overall Rating | 700 | 71 | 242 | 70 |
| Mean of All Ratings | 705 | 68 | 242 | 69 |

*Decimals omitted.

TABLE 12

*Numbers of Detectives Whose Job Performance Was Rated by 1, 2, 3, & 4 Raters*

| Number of Raters | Number of Ratees |
|---|---|
| 1 | 257 |
| 2 | 187 |
| 3 | 39 |
| 4 | 2 |
| | 485 |

mates for these dimensions city by city shows that they were rated very poorly in almost all cities with the possible exception of Minneapolis and Cincinnati.

Again, we decided to discard from further analyses the ratings of lowest quality and to recompute overall reliability estimates across all cities for those sets of data that were most reliable. The following decisions were made:

- Ratings for Cincinnati detectives on dimensions G and L were discarded.
- Ratings for Des Moines detectives on dimensions A, B, F, and L were discarded.
- Ratings for Miami detectives on dimensions E, F, G, H, I, K, and L were discarded.
- Ratings for Portland detectives on dimensions B, C, E, G, I, and J were discarded.
- Ratings for San Diego detectives on dimensions G, H, J, K, and L were discarded.
- Ratings for Washington detectives on dimensions E, F, G, I, J, and K were discarded.

New reliability estimates were then recomputed across all cities for the remaining information. Again, these new estimates were computed for only those detectives for whom test information was available. Table 14 shows the final reliability estimates for tested detectives for each of the dimensions (with data discarded as specified above) across all cities. In most instances, these reliabilities are acceptable, but again drastic reductions in sample size were necessary for many of the dimensions—especially for dimensions E, G, I, J, K, and L. And, in fact, the reliabilities still are considerably lower than desirable for Dimension C (Interrogation), Dimension G (Appearing in Court), Dimension H (Public Relations), Dimension I (Juveniles), and Dimension L (Integrity).

c. *Sergeant ratings.* Table 15 shows the pattern of multiple ratings obtained for sergeants. Examination of reliabilities according to numbers of raters showed that sergeants with three ratings in the cities of Portland, Miami, Washington, and San Diego were rated with especially low reliability. Thus, ratings for a total of 12 sergeants were discarded from further analyses: three from Portland, three from Miami, five from Washington, and one from San Diego. A total of 12 sergeants was tested and rated in Tucson and Albuquerque, but they received only one rating each and could not, therefore,

TABLE 13

*Interrater Reliability Estimates* for Job Performance Ratings for All Detectives (N = 483**)*

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=187) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Investigating a Crime | 61 | 66 | 67 |
| B | Arrest, Search, Seizure | 49 | 57 | 55 |
| C | Interrogation | 46 | 52 | 52 |
| D | Investigating a Case | 54 | 59 | 60 |
| E | Developing Informants | 42 | 33 | 48 |
| F | Report Writing | 30 | 28 | 35 |
| G | Appearing in Court | 27 | 16 | 33 |
| H | Public Relations | 23 | 24 | 28 |
| I | Juveniles | 17 | 20 | 22 |
| J | Cooperation | 43 | 33 | 49 |
| K | Dedication | 38 | 39 | 44 |
| L | Integrity | 21 | 26 | 26 |
| M | Overall Rating | 56 | 61 | 62 |
| | Sum of Ratings | 55 | 59 | 61 |
| | Mean Number of Raters (K) | 1.55 | 2.00 | .. |

*Decimals omitted.

**Two detectives with unreliable multiple ratings have been deleted from this analysis.

TABLE 14

*Interrater Reliability Estimates* for Retained Job Performance Ratings for All Tested Detectives*

| Dimension | Overall Estimate | | Estimate Based on Ratees with Two Ratings Only | |
|---|---|---|---|---|
| | N | r | N | r |
| A Investigating a Crime | 377 | 65 | 140 | 74 |
| B Arrest, Search, Seizure | 334 | 56 | 127 | 64 |
| C Interrogation | 361 | 41 | 142 | 46 |
| D Investigating a Case | 411 | 53 | 156 | 60 |
| E Developing Informants | 209 | 60 | 79 | 58 |
| F Report Writing | 233 | 50 | 76 | 65 |
| G Appearing in Court | 104 | 30 | 48 | 19 |
| H Public Relations | 297 | 32 | 108 | 32 |
| I Juveniles | 178 | 37 | 66 | 44 |
| J Cooperation | 219 | 47 | 99 | 38 |
| K Dedication | 205 | 52 | 78 | 59 |
| L Integrity | 214 | 43 | 75 | 53 |
| M Overall Rating | 415 | 54 | 156 | 63 |
| Mean of All Ratings | 415 | 62 | 156 | 68 |

*Decimals omitted.

be included separately by city in these reliability analyses.

TABLE 15

*Numbers of Sergeants Whose Job Performance Was Rated by 1, 2, 3, Etc. Raters*

| Number of Raters | Number of Ratees |
|---|---|
| 1 | 176 |
| 2 | 178 |
| 3 | 21 |
| 4 | 8 |
| 5 | 4 |
| 6 | 3 |
| | 390 |

Reliability estimates were computed for each of the sergeant dimensions for all cities together and separately for each of the cities. Tables showing separate estimates for each city are given in Appendix B. Table 16 shows the reliability estimates for sergeant dimension ratings for all sergeants taken together. Though none of the estimates is abysmally low, neither are they outstandingly high. With the exception of a few dimensions here and there, only the information from Portland seemed almost uniformly to show low quality ratings for sergeants. Thus, less drastic "surgery" was required in discarding data for sergeants than was the case for patrol officers and detectives. The following decisions were made for discarding sergeants' ratings:

- Ratings for Cincinnati sergeants on dimension D were discarded.
- Ratings for Portland sergeants on dimensions A, B, D, E, G, and H were discarded.
- Ratings for San Diego sergeants on dimension B were discarded.

New reliability estimates were then recomputed across all cities for the remaining information. Again, these new estimates were computed for only those sergeants for whom test information was available. Table 17 shows the final reliability estimates for tested sergeants for each of the dimensions (with data discarded as specified above) across all cities. Surprisingly, the estimates shown in Table 17 are about the same size as those shown in Table 16. Most of the reliabilities are marginally acceptable. None is so low as to rule out continued study of the dimensionality of the performance ratings, though dimensions A and E approach levels of uselessness in their reliabilities.

d. *Middle Command personnel.* For middle command ratings, the only necessary deletion of information was for two persons with four ratings each from Washington. Of the remaining 215 middle command officers, 65 were in departments where no one received more than one rating from a senior officer (no multiple ratings for middle level personnel were obtained from Cincinnati, San Diego, Tucson, or Albuquerque); thus, reliability estimates are available separately for only five cities. Tables showing separate estimates for these five cities are given in Appendix B. Table 18 shows the reliability estimates

for middle command dimension ratings for all middle level personnel taken together. With the exception of dimension C (Scheduling) and dimension H (Public Relations), the reliabilities are reassuringly high. Even so, cities varied somewhat in the reliabilities of the ratings—sufficiently to suggest the wisdom of again discarding some ratings from further analyses. The following decisions were made:

- Dimension H was unreliably rated in all cities. Therefore, all dimension H ratings were discarded from further analyses.
- Ratings for Des Moines middle level officers on dimensions B, E, F, and G were discarded.
- Ratings for Miami middle level officers on dimensions C and G were discarded.

TABLE 16

*Interrater Reliability Estimates* for Job Performance Ratings for All Sergeants (N = 379**)*

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=178) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | 37 | 38 | 41 |
| B | Coordination and Deployment | 46 | 48 | 50 |
| C | Supervision | 51 | 39 | 55 |
| D | Administration and Inspection | 45 | 49 | 49 |
| E | Decision Making and Initiative | 34 | 37 | 38 |
| F | Training and Planning | 52 | 54 | 57 |
| G | Dedication | 40 | 48 | 44 |
| H | Public Contact | 40 | 49 | 44 |
| I | Overall Rating | 57 | 64 | 61 |
| | Sum of Ratings | 58 | 63 | 62 |
| | Mean Number of Raters (K) | 1.70 | 2.00 | .. |

*Decimals omitted.

**Eleven sergeants with unreliable multiple ratings have been deleted from this analysis.

TABLE 17

*Interrater Reliability Estimates* for Retained Job Performance Ratings for All Tested Sergeants*

| | Dimension | Overall Estimate | | Estimate Based on Ratees with Two Ratings Only | |
|---|---|---|---|---|---|
| | | N | r | N | r |
| A | Concern for Subordinates | 350 | 35 | 165 | 38 |
| B | Coordination and Deployment | 298 | 51 | 141 | 55 |
| C | Supervision | 359 | 48 | 170 | 59 |
| D | Administration and Inspection | 284 | 49 | 125 | 54 |
| E | Decision Making Initiative | 349 | 29 | 165 | 36 |
| F | Training and Planning | 363 | 53 | 172 | 53 |
| G | Dedication | 352 | 44 | 167 | 47 |
| H | Public Contact | 352 | 42 | 167 | 49 |
| I | Overall Rating | 362 | 57 | 172 | 63 |
| | Mean of All Ratings | 365 | 58 | 172 | 64 |

*Decimals omitted.

- Ratings for Minneapolis middle level officers on dimension E were discarded.
- Ratings for Washington middle level officers on dimension I were discarded.

New reliability estimates were then recomputed across all cities for all ratings for middle level officers who also had been tested. Results are shown in Table 19. The reliabilities of middle level officer ratings are acceptable, with the exception of those for dimension C (Scheduling). Ratings on dimension D (Training), dimension E (Supervision), and dimension G (Field Command) are only marginally acceptable.

e. *Summary of reliability results.* Table 20 summarizes the reliability estimates for those ratings retained for use in additional analyses. The median values are all in the same range and are acceptably high. Ratings of Overall Performance and the Mean of All Ratings Given (which may be viewed as another form of overall performance index) show high reliabilities for all police jobs. Finally, in each job, the four dimensions with highest reliabilities are exactly those dimensions for each job which appear intuitively to represent the most important and most readily observable aspects of police performance. This is particularly true for the detective and sergeant jobs.

TABLE 18

*Interrater Reliability Estimates* for Job Performance Ratings for All Middle Level Officers (N = 215)*

| Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=58) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|
| A Administrative | 56 | 72 | 65 |
| B Communications | 53 | 67 | 63 |
| C Scheduling | 30 | 43 | 39 |
| D Training | 48 | 57 | 58 |
| E Supervision | 45 | 50 | 55 |
| F Subordinate Evaluation | 59 | 68 | 68 |
| G Field Command | 43 | 32 | 53 |
| H Public Relations | 02 | 20 | 03 |
| I Dedication | 45 | 50 | 55 |
| J Overall Performance | 58 | 55 | 67 |
| Sum of Ratings | 64 | 71 | 73 |
| Mean Number of Raters (K) | 1.35 | | |

*Decimals omitted.

TABLE 19

*Interrater Reliability Estimates* for Retained Job Performance Ratings for All Tested Middle Level Officers*

| Dimension | Overall Estimate | | Estimate Based on Ratees with Two Ratings Only | |
|---|---|---|---|---|
| | N | r | N | r |
| A Administrative | 204 | 59 | 56 | 73 |
| B Communications | 186 | 52 | 48 | 62 |
| C Scheduling | 172 | 30 | 46 | 44 |
| D Training | 204 | 43 | 56 | 56 |
| E Supervision | 158 | 48 | 36 | 65 |
| F Subordinate Evaluation | 186 | 56 | 48 | 67 |
| G Field Command | 150 | 42 | 38 | 02 |
| H Public Relations | —deleted— | | —deleted— | |
| I Dedication | 157 | 52 | 40 | 53 |
| J Overall Performance | 204 | 55 | 56 | 54 |
| Mean of All Ratings | 206 | 65 | 56 | 72 |

*Decimals omitted.

TABLE 20

*Median Reliability Estimates, Dimensions With Highest Reliabilities, and Reliabilities for Overall Rating and Mean of Ratings for Retained Performance Ratings*

| | | | |
|---|---|---|---|
| *Patrol Officers* | | *Detectives* | |
| Median Reliability: | .54 | Median Reliability: | .51 |
| Overall Performance: | .71 | Overall Performance: | .54 |
| Mean of All Ratings: | .68 | Mean of All Ratings: | .62 |
| Highest Dimensions | | Highest Dimensions | |
| Commitment: | .64 | Investigating a Crime: | .65 |
| Investigation: | .60 | Developing Informants: | .60 |
| Teamwork: | .59 | Arrest, Search, Seizure: | .56 |
| Crime Prevention: | .57 | Investigating a Case: | .53 |
| *Sergeants* | | *Middle Level Officers* | |
| Median Reliability: | .49 | Median Reliability: | .52 |
| Overall Performance: | .57 | Overall Performance: | .55 |
| Mean of All Ratings: | .58 | Mean of All Ratings: | .65 |
| Highest Dimensions | | Highest Dimensions | |
| Training and Planning: | .53 | Administration: | .59 |
| Coordination and Deployment: | .51 | Subordinate Evaluation: | .56 |
| Administration and Inspection: | .49 | Communications: | .52 |
| Supervision: | .48 | Dedication: | .52 |

3. *Factor analytic studies.* A major reason for carrying out the rather extended surgical procedures described in the preceding pages was to "clean up the performance ratings" sufficiently to allow us to place confidence in results from dimensional analyses of them. We wished to factor analyze the intercorrelation matrices based on performance ratings in order to reduce the necessary number of dimensions at the item analysis stage of our studies. Thus, as described, we discarded performance ratings of low quality in order to maximize the statistical stabilities of correlations between pairs of dimension ratings and the corresponding stabilities of the subsequent factor matrices derived from the correlation tables. Discarding information had the effect, of course, of producing different sample sizes for performance ratings on various dimensions. [For example, for patrol officers, the sample sizes varied from 360 for dimension I (Domestic Disputes) to 702 for dimension F (Report Writing).] For computation, pairwise deletion of information was used so that each resulting matrix contained correlation coefficients based on different sample sizes. Tables showing final sample sizes, means and standard deviations of the ratings, and the correlation matrices between ratings are given in Appendix C for each of the four police jobs. Information in Appendix C includes data for only those officers for whom both test data and performance ratings were available.

Correlations between all ratings and the two overall indices (the Overall Performance Rating and the Mean of All Ratings) were dropped from each of the four correlation matrices. Each remaining matrix was then factored beginning with multiple $R^2$ in the cells and iterating through varimax rotations to a specified number of factors until the communality estimates stabilized. [See Nice, Bent, & Hall (1970) for a detailed description of this factor analysis routine.] Several solutions with varying numbers of factors were examined for each police job before settling on one solution that seemed to portray most meaningfully the dimensionality of each correlation matrix. Results of these analyses are described below and on subsequent pages.

a. *Patrol officers.* Results for two-factor and three-factor solutions for the patrol officer ratings are shown in Table 21. A large general factor involving Crime Prevention, Investigation, Commitment, and Public Safety is shown by both solutions. A second factor, shown in both solutions, is defined by performance dimensions (Appropriate Use of Force, Dealing with the Public, Handling Domestic Disputes) involving direct personal contact between patrol officers and citizens. The third solution is distinctive in that Teamwork stands out as a useful performance dimension. And, conceptually at least, Teamwork is somewhat different from the core behavior represented by either of the first two factors.

These results were used in conjunction with reliability information and knowledge of the sample sizes available for analysis to decide which dimensions would be used for item analysis purposes. We sought to choose a dimension that: (1) could be regarded as a "marker" dimension for the factor on which it was most highly loaded; (2) showed a high reliability estimate in our previous analyses; and (3) was based on as large a

TABLE 21

*Factor Matrices* for Two-Factor and for Three-Factor Solutions for Job Performance Ratings for Patrol Officers*

| | | Two-Factor Solution** | | | Three-Factor Solution** | | | |
|---|---|---|---|---|---|---|---|---|
| | *Dimension* | I | II | $h^2$ | I | II | III | $h^2$ |
| A | Crime Prevention | [72] | 29 | 60 | [78] | 28 | 10 | 69 |
| B | Use of Force | 10 | [75] | 57 | 15 | [78] | −03 | 63 |
| C | Traffic Control | 64 | 38 | 56 | 59 | 35 | 28 | 55 |
| D | Public Safety | [80] | 21 | 68 | [68] | 18 | 43 | 68 |
| E | Investigation | [86] | 21 | 79 | [83] | 19 | 28 | 81 |
| F | Report Writing | 68 | 30 | 55 | 65 | 28 | 24 | 55 |
| G | Ethical Conduct | 54 | 29 | 37 | 50 | 27 | 21 | 37 |
| H | Dealing with the Public | 46 | [70] | 70 | 36 | [66] | 38 | 72 |
| I | Domestic Disputes | 45 | [75] | 77 | 35 | [72] | 36 | 77 |
| J | Commitment | [78] | 24 | 67 | [70] | 21 | 36 | 66 |
| K | Teamwork | 69 | 26 | 54 | 48 | 18 | [69] | 73 |

*Decimals omitted.

**Note: In the non-rotated solutions, Factor I accounted for 87.5% and 83.3% of the common variance respectively in the two-factor and three-factor solutions.

number of subjects as possible. In choosing a dimension to represent Factor I, the following information was considered:

| *Dimension* | *N* | *Reliability* |
|---|---|---|
| Crime Prevention | 618 | .57 |
| Public Safety | 465 | .46 |
| Investigation | 363 | .60 |
| Commitment | 522 | .64 |

In choosing a dimension to represent Factor II, the following information was considered:

| *Dimension* | *N* | *Reliability* |
|---|---|---|
| Use of Force | 626 | .54 |
| Dealing with the Public | 542 | .48 |
| Domestic Disputes | 360 | .53 |

Obviously, Teamwork with an N of 520 and a reliability of .59 was the only dimension available for representing Factor III.

Using the above guidelines, we chose the following dimensions to serve as criterion scales against which to carry out item analyses for patrol officers:

Crime Prevention, hereinafter called: P-CRIME;
Use of Force, hereinafter called: P-FORCE;
Teamwork, hereinafter called: P-TEAMWORK; and,
the Overall Rating, hereinafter called: P-OVERALL.

b. *Detectives.* Table 22 shows results for three- and four-factor solutions for the correlation matrix based on detective performance ratings. A massive general factor consists of the intuitively most important detective dimensions. The second factor is a curious blend of Integrity and Dealing with Juveniles. Factors III and IV are specific to the dimensions of Public Relations and Report Writing, respectively. The major decision here was to select the best possible dimension to serve as the "marker" for the big general factor. Here is the relevant information for that decision:

| *Dimension* | *N* | *Reliability* |
|---|---|---|
| Investigating a Crime | 377 | .65 |
| Arrest, Search, Seizure | 334 | .56 |
| Interrogation | 361 | .41 |
| Investigating a Case | 411 | .53 |
| Developing Informants | 209 | .60 |
| Dedication | 205 | .52 |

For Factor II, the following were relevant:

| *Dimension* | *N* | *Reliability* |
|---|---|---|
| Juveniles | 178 | .37 |
| Integrity | 214 | .43 |

The reliability of Public Relations dimension ratings was only .32, too low to be considered further. However, the reliability of Report Writing dimension ratings was an acceptable .50, based on 233 subjects.

Again, using guidelines of sample size, level of reliability, and factor definition, the following dimensions were chosen to serve as criterion scales against which to carry out item analyses for detectives:

Investigating a Crime, hereinafter called: D-INVESTIGATE;
Integrity, hereinafter called: D-INTEGRITY;

TABLE 22

*Factor Matrices* for Three-Factor and for Four-Factor Solutions for Job Performance Ratings for Detectives*

| | | Three-Factor Solution** | | | | Four-Factor Solution** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Dimension* | I | II | III | $h^2$ | I | II | III | IV | $h^2$ |
| A | Investigating a Crime | [78] | 18 | 30 | 73 | [76] | 17 | 15 | 32 | 73 |
| B | Arrest, Search, Seizure | [70] | 17 | 42 | 69 | [68] | 12 | 26 | 37 | 68 |
| C | Interrogation | [60] | 39 | 23 | 57 | [66] | 22 | 43 | 02 | 67 |
| D | Investigating a Case | [72] | 26 | 31 | 68 | [72] | 19 | 23 | 28 | 68 |
| E | Developing Informants | [63] | 46 | 07 | 62 | [53] | 44 | 14 | 07 | 62 |
| F | Report Writing | 43 | 16 | [57] | 53 | 36 | 18 | 19 | [66] | 64 |
| G | Appearing in Court | ..... | | | Deleted Due to Low N(104) | | | | ..... | |
| H | Public Relations | 24 | 47 | 44 | 47 | 23 | 24 | [70] | 18 | 64 |
| I | Juveniles | 22 | [77] | 13 | 66 | 26 | [63] | 37 | 00 | 60 |
| J | Cooperation | 34 | 43 | [60] | 66 | 36 | 30 | 52 | 38 | 63 |
| K | Dedication | [69] | 36 | 39 | 75 | [67] | 31 | 27 | 36 | 75 |
| L | Integrity | 22 | 57 | 28 | 45 | 17 | [72] | 15 | 30 | 66 |
| | Percent of Variance | 84.9% | 10.0% | 5.2% | .. | 79.4% | 9.9% | 5.5% | 5.3% | .. |

*Decimals omitted.

**Note: In the non-rotated solutions, Factor I accounted for 84.9% and 79.4% of the common variance respectively in the three-factor and four-factor solutions.

Report Writing, hereinafter called: D-REPORTING; and,
Mean of All Ratings, hereinafter called: D-OVERALL.

c. *Sergeants.* Ratings for sergeants yielded an even larger and more pervasive general factor than those for patrol officers and detectives. Even so, a three-factor solution was meaningful. It is shown in Table 23. Using essentially the same bases for decisions here as for the patrol officers and detectives, the following dimensions were chosen to serve as criterion scales against which to carry out item analyses for sergeants:

Training and Planning, hereinafter called: S-TRAINING;
Supervision, hereinafter called: S-SUPERVISION;
Administration and Inspection, hereinafter called: S-INSPECTION; and,
Overall Performance Ratings, hereinafter called: S-OVERALL.

d. *Middle level command officers.* The four-factor solution provided the most clear-cut picture of the dimensionality of the criterion ratings for middle level command personnel. Results are shown in Table 24. Again, the first factor is pervasive, accounting for a very substantial portion of the common variance. The dimensions loading highest on Factor I are similar in that they involve the guidance and training of subordinates. Factor II has highest loadings on dimensions involving departmental administrative and scheduling activities. Dedication is most strongly represented on Factor III. Factor IV has highest loadings on the interactive behaviors of Communications and Supervision. Evidence related to our final choices of dimensions for item analysis purposes is shown in the following:

| | *Dimension* | *N* | *Reliability* |
|---|---|---|---|
| Factor I: | Training | 204 | .43 |
| | Subordinate Evaluation | 186 | .56 |
| Factor II: | Administrative | 204 | .59 |
| | Scheduling | 172 | .30 |
| Factor IV: | Communications | 186 | .52 |
| | Supervision | 158 | .48 |

Based on the above information, the following dimensions were chosen to serve as criterion scales against which to carry out item analyses for middle level officers:

Administrative, hereinafter called: C-ADMINISTRATIVE;
Communications, hereinafter called: C-COMMUNICATIONS;
Subordinate Evaluation, hereinafter called: C-PERSONNEL;
Dedication, hereinafter called: C-DEDICATION; and,
Mean of All Ratings, hereinafter called: C-OVERALL.

e. *Summarizing comment.* The most obvious feature of the factor analytic results for all four police jobs is the presence of massive general factors in all four sets of performance ratings. It is apparent that our efforts to overcome halo error through painstaking development of behaviorally anchored job performance scales met with

TABLE 23

*Factor Matrix* Showing Three-Factor Solution for Job Performance Ratings for Sergeants*

| | | Three-Factor Solution** | | | |
|---|---|---|---|---|---|
| | *Dimension* | I | II | III | $h^2$ |
| A | Concern for Subordinates | [56] | 44 | 10 | 52 |
| B | Coordination and Deployment | 34 | [55] | 42 | 59 |
| C | Supervision | 27 | [64] | 44 | 68 |
| D | Administration and Inspection | 29 | 35 | [66] | 64 |
| E | Decision Making and Initiative | 36 | [57] | 25 | 52 |
| F | Training and Planning | [56] | 27 | 47 | 61 |
| G | Dedication | [59] | 31 | 29 | 53 |
| H | Public Contact | [65] | 22 | 30 | 56 |

*Decimals omitted.

**Note: In the non-rotated solution, Factor I accounted for 88.9% of the common variance.

TABLE 24

*Factor Matrix* Showing Four-Factor Solution for Job Performance Ratings for Middle Level Command Officers*

| | | Four-Factor Solution** | | | | |
|---|---|---|---|---|---|---|
| | *Dimension* | I | II | III | IV | $h^2$ |
| A | Administrative | 39 | [60] | 18 | 22 | 60 |
| B | Communications | 45 | 29 | 21 | [51] | 61 |
| C | Scheduling | 20 | [61] | 36 | 26 | 62 |
| D | Training | [62] | 31 | 25 | 18 | 57 |
| E | Supervision | 25 | 36 | 31 | [48] | 52 |
| F | Subordinate Evaluation | [69] | 21 | 24 | 27 | 65 |
| G | Field Command | 47 | 25 | 45 | 21 | 54 |
| I | Dedication | 25 | 24 | [51] | 19 | 42 |

*Decimals omitted.

**Note: In the non-rotated solution, Factor I accounted for 86.1% of the common variance.

little success. Most raters apparently formed overall impressions about the relative effectiveness of the officers they were rating and allowed these global impressions to color their ratings on all the other performance rating scales.

A major problem in obtaining job performance ratings for police officers is that superior officers usually have only a limited opportunity to make firsthand, on-the-job observations of how their subordinates are going about their jobs. They rely most often upon reports from the field, and these may suffer from all sorts of problems ranging from incompleteness to actual distortion, either unintentional or intentional. Perhaps peer ratings would have yielded more distinct behavioral descriptions of job performance. Certainly they would have been based more fully on actual firsthand observations of police job performance.

Unfortunately, peer evaluations could not be obtained in these studies. We found that most departments were most comfortable with the traditional concept of superiors rating subordinates, and they felt very uneasy about the notion of co-workers rating each others' job performance. Landy and Goodin (1974) imply that peer evaluation may become more common in the years ahead as the trend toward greater personal professionalism continues to develop in law enforcement agencies. We hope this may be true. Our experience suggests that the superior ratings we obtained were heavily influenced by global perceptions of job effectiveness. We believe peer perceptions of co-workers' job activities in police work will almost always be based on broader observational opportunities. They should, therefore, be much less overwhelmed by global judgments of job effectiveness than our ratings were.

In spite of the large general factors shown by correlation tables based on each of the four sets of ratings, the

factor solutions also showed fairly meaningful additional factors. In each instance, the general factor seemed to be comprised mostly of performance dimensions possessing intuitive appeal as the most salient for the particular job being rated. Thus, for patrol officers, Crime Prevention, Public Safety, Investigation, and Commitment were heavily represented on Factor I. For detectives, Investigating a Crime; Arrest, Search, and Seizure; Interrogation; Investigating a Case; Developing Informants; and Dedication were heavily represented on Factor I. For sergeants, Concern for Subordinates, Training and Planning, Dedication, and Public Contact were heavily represented on Factor I. For middle level officers, the first two rotated factors bore high loadings from dimensions of Training, Subordinate Evaluation, Administration, and Scheduling. The lesser factors for each of the four solutions add further interpretive meaning to the above highly salient activities. Recognizing that these lesser factors are weakly defined and to a large degree indistinct from the general factor, we decided nonetheless to attempt to develop separate prediction keys for them. Table 25 summarizes the performance ratings we finally decided upon for use in the item analyses described in the next section.

## G. Item Analysis and Development of Predictor Scales

1. *General overview of analyses.* The tests and inventories administered to participating police officers included the following:

- EAS Tests 1, 4, 7, and 10
- Biographical History Information: 149 items
- Opinion and Self-Description Inventories: 473 items
- Likes and Dislikes Inventories:
  205 items for patrol officers
  213 items for detectives
  195 items for supervisory officers (sergeants and middle level command)
- Situational Judgment Inventories:
  86 items for patrol officers
  90 items for detectives
  97 items for sergeants
  88 items for middle level officers

Results obtained for the EAS tests are given in a later section. The present section describes the method of analysis used to develop scoring keys likely to be predictive of the various criterion dimensions mentioned in Table 25 of the immediately preceding section. The pools of items chosen for item analysis included those listed above with one further modification. In examining the items of the Biographical History Inventory, we noted that two distinctly different types of items could be

TABLE 25

*Dimensions Chosen for Item Analyses; Reliabilities; and Numbers of Tested Officers for Whom Ratings Were Available*

| | Criterion Scale Designation | Number of Tested Officers Rated | Reliability of Ratings |
|---|---|---|---|
| *Patrol Officers:* | P-CRIME | 618 | .57 |
| | P-FORCE | 626 | .54 |
| | P-TEAMWORK | 520 | .59 |
| | P-OVERALL | 700 | .71 |
| *Detectives:* | D-INVESTIGATE | 377 | .65 |
| | D-INTEGRITY | 214 | .43 |
| | D-REPORTING | 233 | .50 |
| | D-OVERALL | 415 | .62 |
| *Sergeants:* | S-TRAINING | 363 | .53 |
| | S-SUPERVISION | 359 | .48 |
| | S-INSPECTION | 284 | .49 |
| | S-OVERALL | 362 | .57 |
| *Middle Level Officers:* | C-ADMINISTRATIVE | 204 | .59 |
| | C-PERSONNEL | 186 | .56 |
| | C-DEDICATION | 157 | .52 |
| | C-COMMUNICATIONS | 186 | .52 |
| | C-OVERALL | 206 | .65 |

identified. A first group of items portrayed descriptions of actions actually taken at some time in the past. Examples of such items include:

13. Did you ever accept a full-time job where the salary was less than your previous job?
    1. yes
    2. no
41. What size high high school did you attend?
    1. fewer than 100 students
    2. 100–499 students
    3. 500–999 students
    4. 1,000–2,000 students
    5. more than 2,000 students
48. Where did most of your spending money come from during the years you were in high School?
    1. allowance from family
    2. my own earnings
    3. partly allowance, partly earnings
    4. other sources
    5. had no spending money

A total of 56 items of this type was identified.

A second group consisted of a mix of somewhat more subjective items, including current practices, impressions of things in the past, attitudes, expectations, etc. Examples of such items include:

23. How much did your parents emphasize success and getting ahead to you?
    1. constantly
    2. frequently
    3. now and then
    4. rarely
    5. never
71. Which of the following is most likely to make you feel more uncomfortable or unhappy?
    1. having a friend not speak to you
    2. making a mistake in your work
    3. being laughed at when some circumstance makes you look silly
    4. having to introduce yourself to someone you don't know
119. Do you consider yourself as:
    1. nervous
    2. fairly tense
    3. fairly relaxed except when the job tension builds up
    4. fairly relaxed
    5. always relaxed

A total of 87 items of this type was identified. Since these two types of items are distinctly different, we decided to analyze them as separate item pools in addition to considering all of them taken together. The first six items of the inventory, consisting solely of identifying information (such as age, sex, etc.), were deleted for purposes of item analysis.

The following item pools were, therefore, available for developing predictor scoring keys:

| *Item Pool* | *Abbreviation* | *Number* |
|---|---|---|
| Biographical History Information—PAST | BIO-PAST | 56 |
| Biographical History Information—MIXED | BIO-MIX | 87 |
| Total Biographical History Information | BIO-TOT | 143 |
| Opinion and Self-Description Inventories | OSDI | 473 |
| Situational Judgment Inventories | SITNL | 86–97 |
| Inventories of Likes and Dislikes | LD | 195–213 |

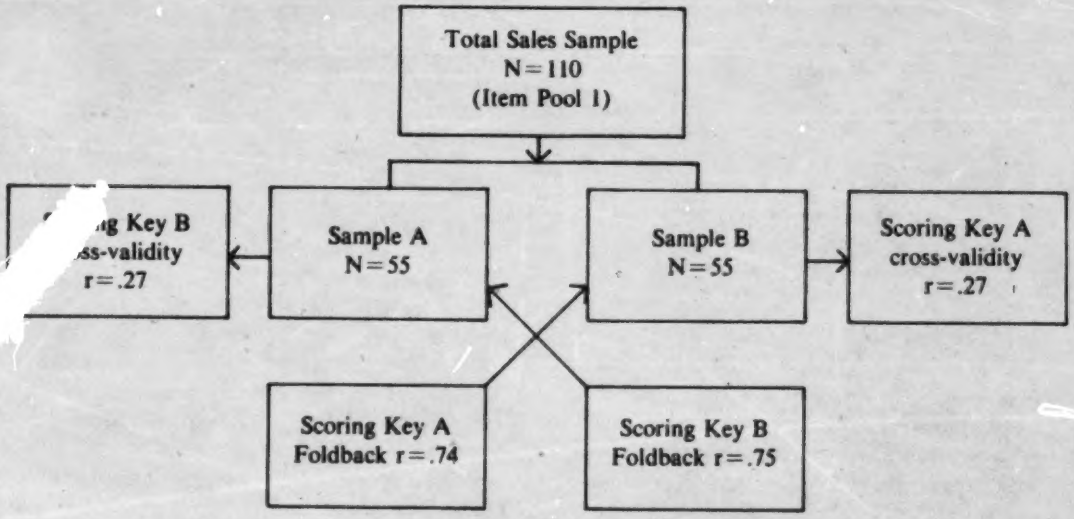
Three major series of analyses were employed in order to develop and select sets of predictor scoring keys for each set of criterion dimensions. The forms of analysis are described below under the headings: "Item analysis and cross-validation," "Modal response keys," and "Factor analyses of predictor-criterion composites."

2. *Item analysis and cross-validation.* The number of separate items available for analysis against the various criterion dimensions varied from 899 for the middle level command inventories to 919 for the detective inventories. When such large numbers of separate item responses are evaluated against some other variable (such as a criterion rating), the likelihood of "discovering" rather large numbers of purely chance relationships is great. Scoring keys based on such item analyses will include not only items with valid and stable relationships with the criterion, but also an unknown number of items with nonstable, chance relationships. The "foldback validity" (i.e., the correlation between a scoring key and a criterion rating computed on the same sample as was used for the item analysis) of such scoring keys is usually high, because the correlation coefficient between key scores and criterion scores reflects both the true, valid relationships and the chance, error-based relationships. When this same scoring key is applied to a new sample of persons, the error component disappears, and the key-criterion correlation is smaller. The validity coefficient is said to "shrink." Such shrinkage is always to be expected whenever a scoring key is developed to give maximum prediction within a sample of persons and is then applied to a different sample of persons.

It is imperative under such circumstances that the methods used for selecting items to be keyed also provide a way of estimating the level of validity (cross-validity) to be expected when the key is applied to subsequent samples of persons. A traditional strategy for providing such estimates has been the strategy of double cross-validation (Katzell, 1951). Double cross-validation typically begins by randomly splitting the available sam-

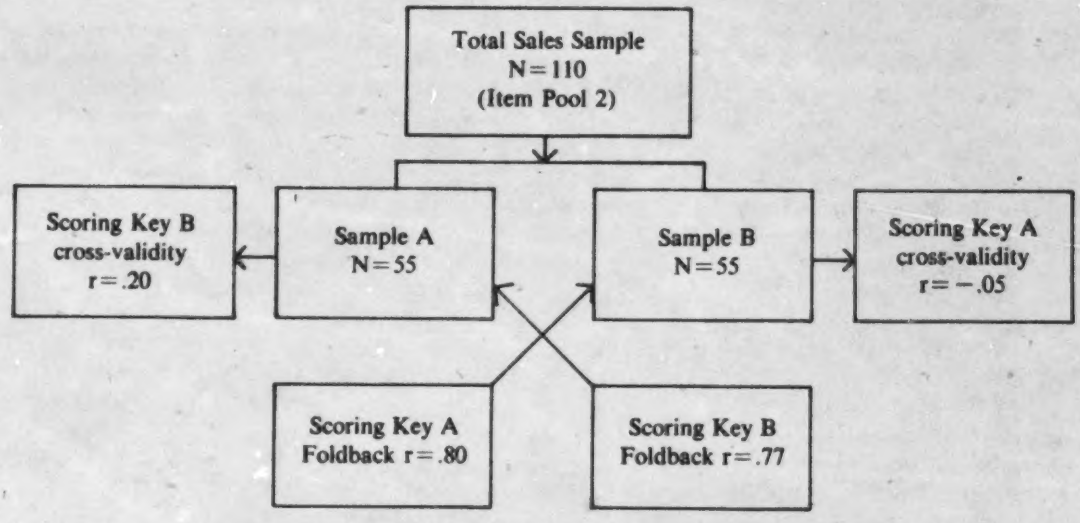
ple of persons into two subsamples equal in size. Scoring keys are then developed separately on each subsample. Each key is then applied to the other subsample in order to estimate its cross-validity.

Here is an example of a double cross-validation design used in developing scoring keys for selecting industrial salesmen (Bownas & Dunnette, 1975):




The above is an example of an ideal outcome. Both keys show the same degree of shrinkage and yield cross-validity estimates of the same magnitude. With such an outcome, we can confidently estimate the degree of shrinkage and the likely usefulness of the keys for use in selection of future candidates for selling.

Unfortunately, double cross-validation designs rarely yield such unambiguous outcomes. A far more frequent pattern is shown below:



Here, one is hard put to decide exactly what to do. Key B might be viewed as showing modest usefulness, but since Key A "washed out" completely, one must remain uncertain about the actual usefulness of any key based on this particular item pool and sample of persons. In effect, a problem with double cross-validation is that the single split of the total sample, even though random, may still yield distinctly different subgroups of persons. In itself, this is not a disadvantage, because one purpose of cross-validation is to determine exactly what may happen to a scoring key's validity when it is applied to different subjects. However, it would be desirable to carry out *several* separate random splits and to compute cross-validity coefficients on many split samples instead of only one as in the double cross-validation design. The median value of the several cross-validity coefficients would provide a more certain estimate of the amount of shrinkage to be expected than the single pair of cross-validity estimates generated by the fortuitous circumstances of just the single random split of the typical double cross-validation approach.

This line of reasoning led us to develop a Monte Carlo procedure for examining the cross-validities of the scoring keys developed in these studies. The procedure begins by splitting the total sample randomly into two subsamples of equal size. A scoring key is developed in one of the samples.* Its foldback validity is determined on that subsample and its cross-validity determined on the other subsample. The subsamples are then recombined to form the total sample which is split again randomly into two equal-sized subsamples, a key developed, foldback and cross-validities computed, subsamples recombined, split anew, a key developed, foldback and cross-validities computed, and so on and so on through as many "realizations" of the process as may be necessary to achieve stability of the distributions of validity coefficients being computed. In our use of this Monte Carlo routine, we find that 15 realizations are sufficient to yield a stable estimate of the median cross-validity of a series of scoring keys developed according to specified criterion limits. The numbers of responses receiving scoring weights depends, of course, on the mean criterion score limits chosen ahead of time. If only a small deviation (e.g., 49.0 or 51.0) from the overall mean of 50.0 is required for a response to be weighted, relatively many responses will receive scoring weights. However, if a larger difference (e.g., 47.0 or 53.0) is required for a response to be weighted, relatively few responses will receive scoring weights. Keys differing in length also will differ in their reliabilities and these, in turn, affect the levels of shrinkage shown on cross-validation. The median cross-validities for keys developed from different criterion score limits and for different item pools are compared. The particular set of limits yielding the least shrinkage is then used in the total sample to fix scoring weights for the item pool that is being analyzed. Fixing final scoring weights on the basis of the *total* sample statistics is best because of increased statistical stability gained from the larger sample size.

Since our Monte Carlo procedure examines the stability of keys based on many random sample splits instead of just one, it is especially well adapted for developing keys in the type of situation we faced in this project. Police officers from many different cities were tested and rated. A single random split might not represent the special features of all cities equally; by using the Monte Carlo procedure, with its multiple realizations, scoring keys made up of items showing systematic validity differences from city to city would tend *not* to yield good cross-validities. Such keys would tend to be dropped from further consideration. Thus, as mentioned, the procedure has the special advantage of discovering scoring keys that are likely to generalize across different settings and different subjects in their levels of validity.

Through the median cross-validities derived from the Monte Carlo routine constitute good (but probably conservative) estimates of the expected usefulness of our scoring keys, the fact of examining several possible keys and choosing those few with the highest median cross-validities could still be argued to involve some element of optimization (that is, some distributions of cross-validities may be high simply by chance and could, therefore, occasionally yield partially spurious results). We decided to attack this potential problem by constructing four special criterion scales, designed with the same means and standard deviations as the OVERALL performance scales for each of the four police job areas. For each job area, the distribution of scores for ratees on the OVERALL criterion dimension were reassigned *ran-*

*The weights for the scoring key are determined as follows. First, the criterion dimension ratings are converted to a standard score distribution with Mean = 50 and Standard Deviation = 10. All response options are then examined to determine the proportion of persons making each response and to compute the mean criterion scores of persons within each response category. A decision is made prior to each computer run to give positive unit weights to response categories where the mean criterion score exceeds a certain value (e.g., 52.5 or 51.5 or 51.0) and negative unit weights to response categories where the mean criterion score falls below a certain value (e.g., 47.5 or 48.5 or 49.0). Also, a prior constraint is set so that no response category will be weighted if fewer than some small proportion (say 5% or 10%) of persons have given responses in that category. Here are some examples of how scoring weights would be assigned for various response distributions to items with the typical *Like* (L), *Indifferent* (I), or *Dislike* (D) format of our Likes and Dislikes Inventories:

1. N = 200 persons
2. No weight given for categories with fewer than 10% of the total responses.

| Item Examples | N | Mean Criterion Score | Criterion Limits 47.5–52.5 | 48.5–51.5 | 49.0–51.0 |
|---|---|---|---|---|---|
| | | | I | I | |
| L | 100 | 50.5 | NO | NO | |
| I | 50 | 51.3 | WEIGHTS | WEIGHTS | |
| D | 50 | 49.1 | I | I | +1 |
| L | 14 | 46.0 | proportion too small | proportion too small | proportion too small |
| I | 150 | 50.5 | | | |
| D | 36 | 48.9 | | | −1 |
| L | 30 | 53.0 | +1 | +1 | +1 |
| I | 40 | 47.2 | −1 | −1 | −1 |
| D | 130 | 50.5 | | | |
| | | | I | | |
| L | 70 | 52.1 | NO | +1 | +1 |
| I | 50 | 50.2 | WEIGHTS | | −1 |
| D | 80 | 48.6 | I | | |

*domly* to the officers in each of the four samples. The reassignments were accomplished by a computer using a random number source. This process yielded four sets of pseudo-ratings with score distributions identical to those of the OVERALL distributions but composed entirely of random error. By definition, the randomly assigned values of this scale could not be correlated with true performance (since they were randomly defined, they could not be correlated with *anything*). Thus, as we proceeded to carry out our item analyses for each of the item pools against the various criterion dimensions, we also carried out identical analyses against each of the four randomly generated pseudo-criterion dimensions. Our hope, of course, was that the median cross-validities for all scoring keys developed against the RANDOM distributions would be very close to 0. This is because such scoring keys presumably would contain only items with nonstable chance relationships and essentially *no* items with truly valid and stable relationships. Any sharp departures from 0 would be worrisome and might well cause us to question the rationale underlying our Monte Carlo procedures.

3. *Modal response keys.* We have just described the basic methodology we employed for developing and selecting empirically weighted scoring keys based on item analyses against criterion performance ratings. The procedure was used with all available item pools. However, we also reasoned that experienced police officers would show certain modal responses to most items on the Situational Judgment Inventories. In other words, certain actions would be regarded by most officers as "correct" or desirable and certain actions would be regarded as "incorrect" or undesirable. Scoring keys based on the modal response patterns of all the officers who completed these inventories were developed for each of the Situational Judgment Inventories. We hypothesized that officers scoring high on such keys were, in effect, demonstrating that they were more fully aware of the "correct" actions to be taken in different police situations than those scoring low. Thus, scores on such modal response keys should be correlated positively with various facets of job performance as shown by criterion scale ratings.

We reasoned similarly in developing modal response keys for the Likes and Dislikes Inventories. Here, we hypothesized that officers scoring high on such keys were, in effect, demonstrating preferences more similar to police officers in general than officers scoring low on such keys. Thus, here too, scores might be expected to be correlated positively with various facets of rated job performance as shown by criterion scale ratings. Appendix D shows the decision rules that were programmed for the computer so that modal response keys could be generated automatically for the four Situational Judgment Inventories and the three Likes and Dislikes Inventories.

4. *Factor analyses of predictor-criterion composites.* After the best scoring keys had been selected for each of the criterion dimensions within a job area, the keys were scored for all the officers in that job area. Correlation coefficients between all possible pairs of scores (both predictor scales and criterion ratings) were computed and the resulting correlation matrix factor was analyzed. This was done in order to discover scoring keys which showed relatively large amounts of common variance so that such keys could be combined to form presumably more reliable and stable predictors for use in the final forms of the POLICE CAREER INDEX.

5. *Item pools and criterion limits for item analyses.* Computer time required to generate cross-validity estimates by the Monte Carlo routine described previously increases rapidly as the number of items in the item pool increases. In the interest of computer cost efficiency, we

decided, therefore, to divide the OSDI items approximately into thirds for item analysis purposes. We already have described the division of biographical items according to examination of their content. The item pools available for analysis, therefore, included the following for each criterion dimension [The numbers in parentheses are the approximate number* of items in each item pool]: BIO-PAST (56), BIO-MIX (87), BIO-TOT (143), LD (200), SITNL (100), OSDI 1 (165), OSDI 2 (155), OSDI 3 (155). Four different criterion limits were established for the analysis of each item pool. These were 47.5 or 52.5; 48.0 or 52.0; 48.5 or 51.5; and, 49.0 or 51.0 Thus, for each criterion dimension, a total of 32 Monte Carlo runs was carried out, 32 scoring keys developed, and their median cross-validities computed. Including the four RANDOM pseudo-criterion dimensions, a total of 21 criterion dimension distributions was available for these item analyses.

Six hundred seventy-two Monte Carlo runs were made across item pools (8), criterion score limits (4), and criterion dimensions (21). Information developed from these 672 computer runs (i.e., numbers of items per scoring key, foldback coefficients, and median cross-validities) is shown in the 21 tables of Appendix E.

The more general results obtained from these item analyses, results with the modal response keys, and the outcomes of the factor analyses are given below and on subsequent pages for each of the four police functions studied.

a. *Patrol officers.* Correlations between the LD and SITNL modal response keys and each of the criterion dimensions for patrol officers are shown in Table 26. Correlations shown there are small. There is possibility that the LD key might, in combination with other keys, have some predictive usefulness. More is said about this later.

For patrol officers, 160 Monte Carlo runs were made, 32 for each of the five criterion distributions. Complete information for all these runs is shown in Appendix E. The following guidelines were used as a basis for deciding which scoring keys to retain for further study:

- The median cross-validities were first examined. Keys with relatively high values were, of course, among the most likely candidates for retention.
- The number of items meeting the criterion score limits for a given item pool for the RANDOM criterion dimension can be regarded as a rough estimate of the number of items to be expected strictly by chance. Designate this number with the symbol $N_R$. Presumably, the number of items meeting the same criterion dimension (e.g., OVERALL) ought to be larger (since valid and stable items will be included along with those that are due only to chance). Designate this number with the symbol $N_C$. The ratio $N_R/N_C$ is an index showing which scoring keys contain relatively fewer valid, stable, nonchance items. In fact, the magnitudes of median cross-validities for various keys correlated −.59 with the size of the $N_R/N_C$ index. Thus, a second basis for choosing keys was to choose keys corresponding to the lowest $N_R/N_C$ ratios.
- Finally, other things being equal, relatively shorter keys were chosen as opposed to keys with very large numbers of items. This guideline is definitely of lesser importance. Nonetheless, it is based on the desire to retain a reasonable degree of simplicity in the final scoring systems for the CAREER INDEX. Also, keys with many items typically would be derived from less rigorous criterion score limits and seem intuitively to contain fewer really "good" items than keys based on more rigorous criterion score limits. As will be seen, this guideline was relaxed in a few instances, particularly in our examination of keys developed from B10 items for the middle level command officers.

TABLE 26

*Correlation Coefficients Between LD and SITNL Modal Response Keys and Criterion Dimension Ratings for Patrol Officers*

| | Modal Response Key | |
|---|---|---|
| *Criterion Dimension* | LD | SITNL |
| P-CRIME (N = 618) | .16 | .10 |
| P-FORCE (N = 626) | .00 | −.02 |
| P-TEAMWORK (N = 520) | .16 | .05 |
| P-OVERALL (N = 700) | .15 | .10 |
| P-RANDOM (N = 700) | .01 | −.01 |

Table 27 summarizes results for the keys identified as best according to the above guidelines for the patrol officer criterion dimensions. Since several of the keys shown in Table 27 were based on subsets of the total OSDI item pool of 473 items, we decided to carry out additional Monte Carlo runs with all OSDI items. These runs resulted in new OSDI keys for each of the four patrol officer criterion dimensions. Also note that the key with lowest median cross-validity for P-OVERALL is the one based on the Situational Judgment Inventory. The Situational Judgment Inventory presents a difficult issue if used to screen inexperienced applicants. They have had no experience in police work. Thus, the test—

*The actual numbers differed slightly, of course, for the different job functions and according to distributions of responses shown by respondents on the OSDI items.

for applicants—could be criticized as unfair. These considerations led us to drop the SITNL key from further consideration for police officers. The inventory could perhaps still be of some use as a training and evaluation instrument during police academy training, however.

The patrol officer scoring keys finally selected for further study via factor analysis are listed in Table 28. Table 28 shows all relevant statistics including odd-even reliability estimates for each of the keys and statistics obtained for keys based on the same item pool with corresponding criterion score limits against P-RANDOM.

In order to study more fully the dimensionality of the ten scoring keys and the five criterion distributions (including the pseudo-criterion: RANDOM), a 15 X 15 correlation matrix was computed and factor analyzed. A four-factor solution was chosen for interpretation because the Random variable showed low loadings on all four factors. (The Random variable received a high loading on Factor V in the five-factor solution, suggesting that a strictly error factor had emerged at that stage of factoring.) Table 29 shows that factor matrix for the four-factor solution. The factor matrix shows that method variance very nearly overwhelms the correlations between individual predictor scales and corresponding criterion dimension ratings. Factor II is defined most strongly by the four OSDI keys. Factor IV is defined by two of the three Likes and Dislikes keys. Highest loadings on Factor III are from the four criterion scales. Only Factor I seems to define predictor and criterion variance cutting across different inventories and criterion ratings. These results are compelling in suggesting that certain scoring keys be combined. We decided, therefore, to reduce the number of predictor scoring keys for patrol officers from ten to four as follows:

- Since the OSDI (Force) key was the only one selected in analyses against the P-FORCE criterion scale, it was retained as a single scoring key and designated PI.
- The remaining OSDI scoring keys were combined to form a single key designated PII.
- The two Likes and Dislikes keys loading high on Factor IV were combined and designated PIII.
- The three keys loading most highly on Factor I [BIO-TOT (Teamwork), LD (Teamwork), BIO-

TABLE 27

*Initial Scoring Keys Chosen for Patrol Officer Criterion Dimensions: Criterion Score Limits, Numbers of Items, Foldback Coefficients, and Median Cross-Validities*

| | Criterion Score Limits | Number of Items Weighted | Foldback Coefficient | Median Cross-Validity |
|---|---|---|---|---|
| *P-CRIME* | | | | |
| BIOTOT | 48.5–51.5 | 56 | .46 | .16 |
| LD | 48.5–51.5 | 52 | .43 | .19 |
| OSDI 1 | 49.0–51.0 | 31 | .37 | .15 |
| OSDI 2 | 48.5–51.5 | 30 | .34 | .16 |
| *P-FORCE* | | | | |
| OSDI 2 | 49.0–51.0 | 31 | .37 | .15 |
| OSDI 3 | 48.5–51.5 | 30 | .34 | .16 |
| *P-TEAMWORK* | | | | |
| BIOTOT | 48.5–51.5 | 70 | .47 | .21 |
| LD | 48.5–51.5 | 60 | .51 | .23 |
| OSDI 2 | 49.0–51.0 | 42 | .42 | .16 |
| OSDI 3 | 48.5–51.5 | 50 | .33 | .21 |
| *P-OVERALL* | | | | |
| BIOTOT | 48.5–51.5 | 50 | .45 | .23 |
| LD | 48.5–51.5 | 46 | .43 | .22 |
| SITNL | 48.5–51.5 | 64 | .54 | .17 |
| OSDI 1 | 49.0–51.0 | 44 | .39 | .23 |
| OSDI 2 | 49.0–51.0 | 47 | .37 | .23 |
| OSDI 3 | 48.5–51.5 | 43 | .29 | .19 |
| *P-RANDOM* | [NO KEYS SHOWED ACCEPTABLE LEVELS OF CROSS-VALIDITY] | | | |

TABLE 28

*Scoring Keys for Patrol Officer Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against P-Random*

| | Criterion Score Limits | Number of Items Weighted | Odd-Even Reliability | Foldback Coefficient | Median Cross-Validity | P-Random Number of Items | P-Random Foldback Coefficient | P-Random Median Cross-Validity |
|---|---|---|---|---|---|---|---|---|
| *P-CRIME* | | | | | | | | |
| BIOTOT (Crime) | 48.5–51.5 | 56 | .55 | .46 | .16 | 30 | .37 | .02 |
| LD (Crime) | 48.5–51.5 | 52 | .63 | .43 | .19 | 23 | .36 | .03 |
| OSDI (Crime) | 48.5–51.5 | 39 | .71 | .39 | .17 | 14 | .27 | .01 |
| *P-FORCE* | | | | | | | | |
| OSDI (Force) | 49.0–51.0 | 71 | .78 | .39 | .19 | 14 | .27 | .01 |
| *P-TEAMWORK* | | | | | | | | |
| BIOTOT (Teamwork) | 48.5–51.5 | 70 | .68 | .47 | .21 | 30 | .37 | .02 |
| LD (Teamwork) | 48.5–51.5 | 60 | .54 | .51 | .23 | 23 | .36 | .03 |
| OSDI (Teamwork) | 48.5–51.5 | 50 | .80 | .33 | .21 | 12 | .28 | .03 |
| *P-OVERALL* | | | | | | | | |
| BIOTOT (Overall) | 48.5–51.5 | 50 | .70 | .45 | .23 | 30 | .37 | .02 |
| LD (Overall) | 48.5–51.5 | 46 | .60 | .43 | .22 | 23 | .36 | .03 |
| OSDI (Overall) | 48.5–51.5 | 86 | .90 | .38 | .24 | 41 | .31 | .00 |
| Median Values | | 52 | .68 | .43 | .21 | 23 | .36 | .02 |

TABLE 29

*Factor Matrix for Four-Factor Solution* of Scoring Key-Criterion Dimension Matrix for Patrol Officers*

| *Scoring Keys* | I | II | III | IV | $h^2$ |
|---|---|---|---|---|---|
| BIOTOT (Crime) | [-59] | -10 | 14 | 42 | 55 |
| LD (Crime) | -33 | -27 | 16 | [75] | 77 |
| OSDI (Crime) | -23 | [-82] | 16 | 28 | 83 |
| OSDI (Force) | 12 | [-86] | 27 | -01 | 82 |
| BIOTOT (Teamwork) | [-83] | -19 | 03 | -16 | 74 |
| LD (Teamwork) | [-71] | -05 | 10 | 19 | 55 |
| OSDI (Teamwork) | -31 | [-81] | -05 | 05 | 75 |
| BIOTOT (Overall) | [-74] | -30 | 04 | 26 | 70 |
| LD (Overall) | -40 | -23 | 14 | [72] | 74 |
| OSDI (Overall) | -23 | [-91] | 12 | 13 | 91 |
| *Criterion Dimensions* | | | | | |
| P-CRIME | -31 | -06 | [75] | 30 | 75 |
| P-FORCE | 20 | -21 | [75] | -06 | 64 |
| P-TEAMWORK | [-64] | -11 | 51 | -16 | 70 |
| P-OVERALL | -43 | -14 | [75] | 17 | 80 |
| P-RANDOM | 10 | 01 | -01 | 35 | 13 |
| Total Variance | 3.29 | 3.22 | 2.13 | 1.74 | 10.38 |

*Decimals omitted.

TOT (Overall)] were combined and designated PIV.

Reliabilities have been computed for these four final keys. Their median cross-validities have also been estimated by the mathematical procedure described in a technical note in Appendix F. Finally, the keys have been correlated with each of the criterion rating dimensions. Results of all these computations are shown in Table 30. Intercorrelations between the four predictor scales are shown in Table 31.

b. *Detectives.* Correlations between the LD and SITNL modal response keys and each of the criterion measures for detectives are shown in Table 32. Obviously, the correlations are too small to be of any use for predicting performance effectiveness for detectives.

One hundred sixty Monte Carlo runs also were made for detectives, 32 for each of the five criterion distributions. Again, complete information for these runs is given in Appendix E. The same steps were followed in selecting a preliminary set of scoring keys for detectives as those already described in detail for patrol officers. The detective scoring keys finally selected for further study via factor analysis are shown in Table 33 along with all relevant statistics for each of the keys.

In order to study more fully the dimensionality of the seven scoring keys and the five criterion distributions (including the pseudo-criterion: RANDOM), a 12 X 12 correlation matrix was computed and factor analyzed. A three-factor solution was chosen for interpretation, again because the Random variable showed low loadings on all three factors. Table 34 shows the factor matrix for the three-factor solution.

Factor I is a large general effectiveness factor defined by high loadings from three of the Biographical scoring keys [BIO-TOT (Investigate), BIO-MIX (Reporting), BIO-MIX (Overall)] and the three corresponding criterion dimensions (INVESTIGATE, REPORTING, OVERALL). Interestingly, Factor II appears to be defined by the criterion ratings on INTEGRITY and the two scoring keys selected against the Integrity criterion. The OSDI (Reporting) key shows a high loading on Factor III, but its low reliability (.30) seems to rule it out as a useful key for use in the Detective CAREER INDEX. Finally, the LD (Reporting) key shows only marginal communality (.42) and also quite low reliability (.41). Based on these results, certain keys were discarded and others combined so that the number of predic-

TABLE 30

*Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Four Scoring Keys (Predictor Scales) for the Patrol Officer Career Index**

| Predictor Scale | N | Scale Statistics: Odd-Even Reliability | Scale Statistics: Estimated Cross-Validity | Patrol Officer Ratings: CRIME | FORCE | TEAMWORK | OVERALL | RANDOM |
|---|---|---|---|---|---|---|---|---|
| PI | 502 | 78 | 19 | 18 | 40 | 13 | 24 | 01 |
| PII | 452 | 93 | 26 | 29 | 14 | 32 | 37 | 03 |
| PIII | 528 | 76 | 23 | 41 | 11 | 28 | 42 | 04 |
| PIV | 372 | 82 | 28 | 37 | 06 | 54 | 48 | 03 |

*Decimals omitted.

TABLE 31

*Intercorrelations* Between Predictor Scales Comprising the Patrol Officer Career Index*

| | PI | PII | PIII | PIV |
|---|---|---|---|---|
| PI | .. | 72 | 22 | 16 |
| PII | 72 | .. | 45 | 43 |
| PIII | 22 | 45 | .. | 49 |
| PIV | 16 | 43 | 49 | .. |

*Decimals omitted.

TABLE 32

*Correlation Coefficients Between LD and SITNL Modal Response Keys and Criterion Dimension Ratings for Detectives*

| Criterion Dimension | Modal Response Key: LD | SITNL |
|---|---|---|
| D-INVESTIGATE (N = 377) | .01 | .10 |
| D-INTEGRITY (N = 214) | .00 | −.12 |
| D-REPORTING (N = 233) | .04 | −.10 |
| D-OVERALL (N = 415) | .06 | .04 |
| D-RANDOM (N = 415) | .04 | −.02 |

TABLE 33

*Scoring Keys for Detective Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against D-Random*

| | Criterion Score Limits | Number of Items Weighted | Odd-Even Reliability | Foldback Coefficient | Median Cross-Validity | D-Random Number of Items | D-Random Foldback Coefficient | D-Random Median Cross-Validity |
|---|---|---|---|---|---|---|---|---|
| *D-INVESTIGATE* | | | | | | | | |
| BIOTOT (Investigate) | 48.5–51.5 | 69 | .63 | .57 | .20 | 53 | .48 | .00 |
| *D-INTEGRITY* | | | | | | | | |
| BIOPAST (Integrity) | 48.5–51.5 | 62 | .55 | .56 | .19 | 27 | .36 | −.02 |
| OSDI (Integrity) | 47.5–52.5 | 17 | .47 | .42 | .13 | 2 | .19 | .03 |
| *D-REPORTING* | | | | | | | | |
| BIOMIX (Reporting) | 48.5–51.5 | 52 | .60 | .62 | .25 | 24 | .40 | .02 |
| LD (Reporting) | 47.5–52.5 | 30 | .41 | .56 | .15 | 12 | .36 | −.03 |
| OSDI (Reporting) | 48.0–52.0 | 29 | .30 | .56 | .17 | 10 | .33 | −.01 |
| *D-OVERALL* | | | | | | | | |
| BIOMIX (Overall) | 49.0–51.0 | 62 | .66 | .58 | .26 | 49 | .48 | −.01 |
| Median Values | | 52 | .55 | .56 | .19 | 24 | .26 | −.01 |

TABLE 34

*Factor Matrix for Three-Factor Solution* of Scoring Key-Criterion Dimension Matrix for Detectives*

| *Scoring Keys* | I | II | III | $h^2$ |
|---|---|---|---|---|
| BIOTOT (Investigate) | [-82] | 06 | −04 | 68 |
| BIOPAST (Integrity) | −05 | [-76] | 13 | 59 |
| OSDI (Integrity) | 13 | [-62] | −38 | 54 |
| BIOMIX (Reporting) | [-70] | 08 | −40 | 65 |
| LD (Reporting) | −45 | −17 | −44 | 42 |
| OSDI (Reporting) | −21 | −13 | [-81] | 71 |
| BIOMIX (Overall) | [-83] | 05 | −12 | 70 |
| *Criterion Dimensions* | | | | |
| D-INVESTIGATE | [-79] | −24 | 13 | 70 |
| D-INTEGRITY | −32 | [-85] | 03 | 82 |
| D-REPORTING | [-67] | −31 | −49 | 79 |
| D-OVERALL | [-81] | −34 | 04 | 78 |
| D-RANDOM | 08 | 18 | −36 | 17 |
| Total Variance | 3.96 | 2.03 | 1.56 | 7.55 |

*Decimals omitted.

tor scoring keys for detectives was reduced from seven to two as follows:

- The three BIO keys loading high on Factor I were combined and designated DI.
- The two scoring keys loading high on Factor II were combined and designated DII.

Reliabilities of predictor scales DI and DII, their estimated cross-validities, and their correlations with the criterion dimension ratings are shown in Table 35. The two predictor scales, DI and DII, are virtually independent. They correlate only .02 with each other.

c. *Sergeants.* Correlations between the LD and

TABLE 35

*Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Two Scoring Keys (Predictor Scales) for the Detective Career Index**

| Predictor Scale | N | Scale Statistics: Odd-Even Reliability | Scale Statistics: Estimated Cross-Validity | Detective Ratings: INVESTIGATE | INTEGRITY | REPORTING | OVERALL | RANDOM |
|---|---|---|---|---|---|---|---|---|
| DI | 245 | 82 | 30 | 52 | 23 | 58 | 59 | 00 |
| DII | 263 | 53 | 21 | 11 | 62 | 26 | 17 | −09 |

*Decimals omitted.

SITNL modal response keys and each of the criterion measures for sergeants are shown in Table 36. The correlations shown there are too small to be of any practical use in promotion decisions related to sergeants.

TABLE 36

*Correlation Coefficients Between LD and SITNL Modal Response Keys and Criterion Dimension Ratings for Sergeants*

| Criterion Dimension | Modal Response Key: LD | SITNL |
|---|---|---|
| S-TRAINING (N = 363) | .04 | .06 |
| S-SUPERVISION (N = 359) | .06 | −.03 |
| S-INSPECTION (N = 284) | .04 | −.10 |
| S-OVERALL (N = 362) | .13 | −.03 |
| S-RANDOM (N = 362) | .02 | −.02 |

One hundred sixty Monte Carlo runs also were made for sergeants, 32 for each of the five criterion distributions. Complete information for these runs is given in Appendix E. The same steps were followed in selecting a preliminary set of scoring keys for sergeants as those already described in detail for patrol officers and detectives. The sergeant scoring keys finally selected for further study via factor analysis are shown in Table 37 along with all relevant statistics for each of the keys. Information in Table 37 shows that the yield of potentially useful keys was considerably poorer for sergeants than it was for both patrol officers and detectives. The reason for this is not known.

In order to study more fully the dimensionality of the five scoring keys and the five criterion distributions (including the pseudo-criterion: RANDOM), a 10 × 10 correlation matrix was computed and factor analyzed. A two-factor solution was chosen for interpretation, again because the Random variable showed low loadings on

TABLE 37

*Scoring Keys for Sergeant Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against S-Random*

| | Criterion Score Limits | Number of Items Weighted | Odd-Even Reliability | Foldback Coefficient | Median Cross-Validity | S-Random: Number of Items | S-Random: Foldback Coefficient | S-Random: Median Cross-Validity |
|---|---|---|---|---|---|---|---|---|
| *S-TRAINING* | | | | | | | | |
| BIOTOT (Training) | 48.0–52.0 | 48 | .52 | .52 | .19 | 35 | .52 | .05 |
| *S-SUPERVISION* | | | | | | | | |
| SITNL (Supervision) | 48.5–51.5 | 78 | .42 | .61 | .13 | 65 | .58 | −.04 |
| *S-INSPECTION* | | | | | | | | |
| BIOMIX (Inspection) | 47.5–52.5 | 19 | .22 | .51 | .16 | 12 | .34 | .00 |
| *S-OVERALL* | | | | | | | | |
| BIOTOT (Overall) | 48.0–52.0 | 39 | .41 | .58 | .17 | 35 | .52 | .05 |
| SITNL (Overall) | 48.5–51.5 | 84 | .44 | .61 | .17 | 65 | .58 | −.04 |
| Median Values | | 48 | .42 | .58 | .17 | 35 | .52 | .00 |

both factors. Table 38 shows the factor matrix for the two-factor solution. Factor I is a general factor cutting across the criterion dimension ratings and carrying high loadings from the two situational keys. Factor II derives high loadings from the biographical keys and also carries moderate loadings on all the criterion dimensions except SUPERVISION. The 19-item biographical key for INSPECTION, though loaded well on Factor II, has such low reliability (.22) that it is not likely to be useful. Otherwise, the most straightforward way of forming predictor keys was to combine the remaining two biographical keys and the two situational keys to form predictor scales SI and SII. SI denotes the Situational Inventory key; SII denotes the biographical key.

Reliabilities of predictor scales SI and SII, their estimated cross-validities, and their correlations with criterion dimension ratings are shown in Table 39. Fortunately, the reliabilities for predictor scales SI and SII are considerably higher than any of those for the initial keys shown in Table 37. Each of the scales correlates well with all of the criterion ratings. The two scales are, however, relatively independent of each other. The correlation between them is .31.

d. *Middle level command officers.* Correlations between the LD and SITNL modal response keys and each of the criterion measures for middle level command officers are shown in Table 40. The modal response key for the Situational Inventory shows much promise. However, the entire inventory contains nearly 100 items. We, therefore, developed a variation of the Monte Carlo procedure which examines the distribution of correlations between each of the items of a test and the criterion, selects the best N items, and proceeds to calculate the cross-validities over several realizations for tests N items long. In addition to the usual item analyses for middle level command officers, we also conducted the additional Monte Carlo runs designed to select the best subset of modal response scored items from the Situational Inventory. In all, 204 Monte Carlo runs were made for the test and rating information from middle level command personnel. The various scoring keys finally selected for further study via factor analysis are shown in Table 41.

TABLE 38

*Factor Matrix for Two-Factor Solution* of Scoring Key-Criterion Dimension Matrix for Sergeants*

| *Scoring Keys* | I | II | $h^2$ |
|---|---|---|---|
| BIOTOT (Training) | −20 | [−84] | 74 |
| SITNL (Supervision) | [−83] | 19 | 73 |
| BIOMIX (Inspection) | −22 | [−66] | 48 |
| BIOTOT (Overall) | −40 | [−67] | 62 |
| SITNL (Overall) | [−74] | −10 | 55 |
| *Criterion Dimensions* | | | |
| S-TRAINING | [−62] | −49 | 62 |
| S-SUPERVISION | [−80] | −14 | 67 |
| S-INSPECTION | [−69] | −39 | 63 |
| S-OVERALL | [−78] | −39 | 76 |
| S-RANDOM | 04 | −16 | 03 |
| Total Variance | 3.60 | 2.22 | 5.82 |

*Decimals omitted.

TABLE 39

*Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Two Scoring Keys (Predictor Scales) for the Sergeant Career Index**

| Predictor Scale | Scale Statistics N | Odd-Even Reliability | Estimated Cross-Validity | Sergeant Ratings TRAINING | SUPERVISION | INSPECTION | OVERALL | RANDOM |
|---|---|---|---|---|---|---|---|---|
| SI | 271 | 62 | 19 | 45 | 55 | 48 | 59 | 00 |
| SII | 286 | 65 | 22 | 54 | 37 | 40 | 54 | 06 |

*Decimals omitted.

In order to study more fully the dimensionality of the 13 scoring keys and the six criterion distributions (including the pseudo-criterion: Random), a 19 X 19 correlation matrix was computed and factor analyzed. A four-factor solution was chosen for interpretation, again because the Random variable showed low loadings on all four factors. Table 42 shows the factor matrix for the four-factor solution.

Factor I is a large general factor with high loadings from biographical and OSDI scoring keys and two important criterion dimensions (ADMINISTRATION and OVERALL). Of the highly loaded scoring keys, the OSDI keys for COMMUNICATIONS and OVERALL have reliabilities of only .42 and .43, respectively. Factor II is defined almost entirely by its high loadings on the three modal response keys of the Situational Inventory. Factor III is a dedication factor, receiving high loadings from the criterion dimension DEDICATION and from the two keys validated against it. Finally, Factor IV derives its highest loading from the criterion dimension PERSONNEL and high loadings also from two of the three keys validated against it and from two other criterion dimensions.

TABLE 40

*Correlation Coefficients Between LD and SITNL Modal Response Keys and Criterion Dimension Ratings for Middle Level Command Officers*

| | Modal Response Key | |
|---|---|---|
| *Criterion Dimension* | LD | SITNL |
| C-ADMINISTRATIVE (N = 204) | .13 | .25 |
| C-PERSONNEL (N = 186) | .07 | .19 |
| C-DEDICATION (N = 157) | −.11 | .13 |
| C-COMMUNICATIONS (N = 186) | .08 | .14 |
| C-OVERALL (N = 206) | .11 | .28 |
| C-RANDOM (N = 206) | −.04 | .06 |

Based on these results, the following decisions were made for discarding some keys and combining others to form five predictor scales for the middle level command officer CAREER INDEX:

- The OSDI (Administration) and BIO-TOT (Administration) keys were combined to form a scale designated as CI.
- Since the BIO-TOT (Overall) key showed moderate

TABLE 41

*Scoring Keys for Middle Command Criterion Dimensions: Relevant Statistics for Keys and for Similar Keys Against C-Random*

| | | | | | | C-Random | | |
|---|---|---|---|---|---|---|---|---|
| | Criterion Score Limits | Number of Items Weighted | Odd-Even Reliability | Foldback Coefficient | Median Cross-Validity | Number of Items | Foldback Coefficient | Median Cross-Validity |
| *C-ADMINISTRATION* | | | | | | | | |
| BIOTOT (Admin) | 49.0–51.0 | 152 | .81 | .73 | .24 | 136 | .72 | .04 |
| SITNLMODAL (Admin) | .. .. | 44 | .44 | .53 | .27 | 44 | .42 | .05 |
| OSDI (Admin) | 48.0–52.0 | 41 | .63 | .59 | .23 | 19 | .33 | −.14 |
| *C-COMMUNICATIONS* | | | | | | | | |
| OSDI (Comm) | 47.5–52.5 | 20 | .42 | .52 | .17 | 7 | .29 | −.10 |
| *C-PERSONNEL* | | | | | | | | |
| BIOMIX (Pers) | 49.0–51.0 | 74 | .65 | .63 | .19 | 67 | .61 | −.02 |
| SITNLMODAL (Pers) | .. .. | 44 | .49 | .51 | .24 | 44 | .42 | .05 |
| SITNLEMPIRIC (Pers) | 48.0–52.0 | 84 | .59 | .68 | .17 | 62 | .65 | −.07 |
| *C-DEDICATION* | | | | | | | | |
| BIOTOT (Dedic) | 48.0–52.0 | 90 | .75 | .76 | .30 | 69 | .68 | .04 |
| SITNLEMPIRIC (Dedic) | 48.0–52.0 | 87 | .66 | .65 | .17 | 62 | .65 | −.07 |
| *C-OVERALL* | | | | | | | | |
| BIOTOT (Overall) | 49.0–51.0 | 161 | .81 | .71 | .34 | 136 | .72 | .04 |
| SITNLMODAL (Overall) | .. .. | 44 | .49 | .60 | .30 | 44 | .42 | .05 |
| LD (Overall) | 48.0–52.0 | 59 | .82 | .54 | .17 | 42 | .59 | .05 |
| OSDI (Overall) | 47.5–52.5 | 35 | .43 | .60 | .21 | 15 | .35 | −.11 |
| Median Values | | 59 | .63 | .60 | .23 | 44 | .59 | −.01 |

TABLE 42

*Factor Matrix for Four-Factor Solution* of Scoring Key-Criterion Dimension Matrix for Middle Level Command Officers*

| *Scoring Keys* | I | II | III | IV | $h^2$ |
|---|---|---|---|---|---|
| BIOTOT (Admin) | [-72] | -28 | -25 | -20 | 69 |
| SITNLMODAL (Admin) | -21 | [91] | -07 | -09 | 88 |
| OSDI (Admin) | [-83] | 10 | -10 | -09 | 72 |
| OSDI (Comm) | [-74] | -02 | -10 | -14 | 58 |
| BIOMIX (Pers) | -25 | 19 | -29 | [-65] | 60 |
| SITNLMODAL (Pers) | -03 | [86] | -08 | -35 | 87 |
| SITNLEMPIRIC (Pers) | -01 | 22 | -03 | [-81] | 70 |
| BIOTOT (Dedic) | -16 | 12 | [-86] | -17 | 80 |
| SITNLEMPIRIC (Dedic) | -19 | 11 | [-81] | -04 | 70 |
| BIOTOT (Overall) | -53 | 25 | [-39] | [-45] | 70 |
| SITNLMODAL (Overall) | -16 | [88] | -21 | -22 | 89 |
| LD (Overall) | -30 | 14 | -15 | -48 | 36 |
| OSDI (Overall) | [-77] | 03 | -17 | -28 | 70 |
| *Criterion Dimensions* | | | | | |
| C-ADMINISTRATION | [-73] | 26 | -11 | -29 | 69 |
| C-COMMUNICATIONS | -41 | 06 | -11 | [-64] | 59 |
| C-PERSONNEL | -24 | 13 | -11 | [-86] | 82 |
| C-DEDICATION | -17 | 07 | [-85] | -28 | 83 |
| C-OVERALL | -52 | -24 | -31 | [-65] | 85 |
| C-RANDOM | -20 | 18 | 15 | 21 | 14 |
| Total Variance | 3.99 | 2.79 | 2.68 | 3.63 | 13.10 |

*Decimals omitted.

loadings across all factors, it was retained as a separate key and was designated as CII.

- The three modal response keys were combined and designated as CIII.
- The BIO-TOT (Dedic) key showed high reliability (.75) and a high factor loading (.86) on Factor III. Its median estimated cross-validity was .30 as opposed to the lower value of .17 for the other key [SITNLEMPIRIC(Dedic)] loading high on Factor III. Hence, the BIO-TOT (Dedic) key was shown to represent Factor III and was designated as CIV.
- Finally, the two keys loading most highly on Factor IV were combined and designated as CV.

Table 43 summarizes all relevant statistics for the various predictor scales chosen to represent promotional potential in relation to scores on the Middle Level Command PCI. Table 44 shows the correlations between the five command level predictor scales.

e. *Summary of predictor scale development.* The preceding pages describe the steps taken in developing and selecting predictor scales for use in evaluating effectiveness potential for four police job areas. A total of 13

TABLE 43

*Reliabilities, Estimated Cross-Validities, and Correlations With Criterion Dimension Ratings for Five Scoring Keys (Predictor Scales) for the Middle Level Command Career Index**

| | Scale Statistics | | | Middle Command Ratings | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Predictor Scale | N | Odd-Even Reliability | Estimated Cross-Validity | ADMINIS-TRATION | COMMUNI-CATIONS | PERSONNEL | DEDICATION | OVERALL | RANDOM |
| CI | 123 | 83 | 24 | 76 | 42 | 43 | 35 | 65 | 07 |
| CII | 140 | 81 | 34 | 55 | 51 | 52 | 48 | 72 | 01 |
| CIII | 193 | 80 | 32 | 40 | 31 | 36 | 26 | 48 | 05 |
| CIV | 158 | 75 | 30 | 27 | 27 | 31 | 77 | 44 | -08 |
| CV | 131 | 74 | 25 | 37 | 50 | 74 | 38 | 61 | -08 |

*Decimals omitted.

TABLE 44

*Intercorrelations* Between Predictor Scales Comprising the Middle Command Officer Career Index*

| | CI | CII | CIII | CIV | CV |
|---|---|---|---|---|---|
| CI | .. | 79 | 42 | 43 | 50 |
| CII | 79 | .. | 43 | 54 | 68 |
| CIII | 42 | 43 | .. | 28 | 43 |
| CIV | 43 | 54 | 28 | .. | 40 |
| CV | 50 | 68 | 43 | 40 | .. |

*Decimals omitted.

scales has been selected, four for patrol officers, two for detectives, two for sergeants, and five for middle level command officers. Relevant information and summary statistics for these scales are summarized in Table 45.

The reliabilities of these scales are adequate, and their correlations with the criterion rating scales are acceptable for prediction purposes. Most impressive are the near-zero correlations between the predictor scales and the Random distributions. This is good evidence that our Monte Carlo item analysis strategy was successful in identifying and weighting item responses showing stable relationships with criterion ratings, and that predictor scale scores, as expected, show no relationship with randomly generated score distribution. Finally, the cross-validities are of acceptable magnitude. The various predictor scales for each of the four police functions will be equally weighted and combined to yield a single estimate of potential in each police job function. The combined scores undoubtedly will be even more reliable and more valid than any of the median values shown in Table 45. Ordinarily, we would have estimated the overall validities to be expected by using the Monte Carlo routine to compute multiple correlation coefficients and distributions of cross validities for the combined predictor scales. Unfortunately, this could not be carried out here because of the varying sample sizes on which the predictor scale-criterion relationships are based. At any

TABLE 45

*Item Content, Median Reliabilities, Median Correlations Against Criterion and Pseudo-Criterion (Random) Distributions, Median Estimated Cross-Validities, and Median Intercorrelations Between Predictor Scale Comprising Four Career Index Inventories*

| CAREER INDEX | Item Content | Median Reliability | Median Correlation Against Criterion Ratings | Median Correlation with Pseudo-Criterion (RANDOM) Distribution | Median Estimated Cross-Validity | Median Intercorrelation Between Scales |
|---|---|---|---|---|---|---|
| *Patrol Officer* | | | | | | |
| PI | OSDI (Personality) | | | | | |
| PII | OSDI (Personality) | | | | | |
| PIII | LD (Interests) | .80 | .30 | .03 | .24 | .44 |
| PIV | LD & BIO (Interests & Background) | | | | | |
| *Detective* | | | | | | |
| DI | BIO (Background) | | | | | |
| DII | OSDI & BIO (Personality & Background) | .67 | .39 | −.05 | .25 | .02 |
| *Sergeant* | | | | | | |
| SI | SITNL (Situational Judgment) | .63 | .51 | .03 | .20 | .31 |
| SII | BIO (Background) | | | | | |
| *Middle Command* | | | | | | |
| CI | OSDI & BIO (Personality & Background) | | | | | |
| CII | BIO (Background) | | | | | |
| CIII | SITNL (Situational Judgment) | .80 | .44 | .01 | .30 | .43 |
| CIV | BIO (Background) | | | | | |
| CV | SITNL & BIO (Situational Judgment & Background) | | | | | |

rate, the median cross-validities of .24, .25, .20, and .30 should be regarded as minimum estimates of the validities to be expected when the CAREER INDEX inventories are put into actual operational use.

The interested reader may wish to refer to Appendix G. There, we have included a sampling of the items and corresponding response scoring weights for each of the predictor scales summarized in Table 45. Examining these illustrative items should give the reader an intuitive sense of the patterns of personal preferences, attitudes, opinions, and background characteristics related to effective and ineffective performance in patrol officer, detective, sergeant, and middle command functions.

Based on the weights for the few sample items shown there, higher rated patrol officers appear to be thoughtful and analytic, persistent, civic minded, confident, attendant (sensitive) to criticism, were accorded leadership status when young, and are physically active yet relaxed in their manner. Higher rated detectives appear to have been "involved" in high school but *not* highly sociable or affiliative. They are intent on gathering much information before making decisions, and they have no apparent difficulty in taking orders. Higher rated sergeants possess a definite self-perception of success, and they show a strong interest in work, valuing it in its own right instead of as merely a means to an end. Higher rated middle command officers are strongly achievement oriented, like people, emphasize work and effort, were active in school and held leadership posts while there, and probably have had some college education.

## H. Rating, Predictor Scale, and Validity Information According to Cities and Subgroups

1. *Performance ratings.*

a. *Tested versus non-tested officers.* Recall that considerably more patrol officers were rated in the cities of Minneapolis and Portland than the number who finally took the experimental test battery. This was done as a way of forming a crude estimate of the relative representativeness of the level of job performance shown by the tested officers. Table 46 shows the means, standard deviations, and t values for Portland and Minneapolis patrol officers on the patrol officer job performance dimensions. Three of the mean differences shown in Table 46 are sufficiently large to be regarded as statistically stable. However, in Minneapolis the tested officers received ratings slightly lower than the officers who were rated but not tested. Just the opposite was the case in Portland, where the officers who appeared for testing were higher rated on the average than those who were rated but not tested. Since the magnitudes of the differences are not large (in no case exceeding half a standard deviation) and since the differences are in opposite directions, we could come to no firm conclusions about the nature of biasing effects, if any, in our tested samples. Thus, we let the matter remain a mystery and proceeded with the analyses.

b. *Criterion differences between cities.* Differences in the levels of ratings obtained for officers from different cities were examined by comparing mean ratings obtained by all tested white male officers. The comparisons were made among whites and males in order to avoid introducing obvious city x race and city x sex interactions into the mean comparisons. Table 47 shows means and standard deviations on the Overall Criterion rating scales for each of the police functions.* Large differences between the means for particular cities are shown for all functions: Minneapolis vs. Cincinnati, Des Moines, and Portland for patrol officers; Portland and San Diego vs. Minneapolis for detectives; Cincinnati and San Diego vs. Des Moines for sergeants; and Cincinnati, Washington, and Miami vs. Des Moines and Minneapolis for middle command officers. Since the samples from each city were selected differently and since they were not equated for such factors as length of service, age, etc., the mean differences could reflect, to a degree at least, valid differences in job performance. At any rate, we had no way of knowing or estimating the degree to which rating errors (such as leniency) might be operating differentially from city to city and from function to function. Moreover, since our job performance rating scales were anchored with actual incidents illustrative of different levels of performance effectiveness, we had good reason to believe that the various mean differences shown in Table 47 actually do reflect absolute differences in average effectiveness for officers in these different departments. Therefore, we decided to use the rating information as obtained instead of standardizing each city's information according to its own mean and standard deviation.

c. *Criterion differences by race and sex.* Table 48 shows the numbers of men, women, whites, and non-whites contained in our samples according to the four functions. Mean ratings and standard deviations are shown for each of the criterion dimensions separately for race and sex subgroups in Tables 49, 50, 51, and 52.** Non-white sergeants received lower mean ratings than

*Comparisons on *only* the OVERALL scales are shown for the sake of clarity. Comparisons were made on all scales, but the general trend is illustrated accurately by presenting only the information on the Overall scales.

**We have not included information for the thirteen female patrol officers, the nine female sergeants, or the six non-white command officers in Tables 49, 51, and 52 because the samples are so small.

TABLE 46

*Means and Standard Deviations on Criterion Dimensions and T Values for Minneapolis and Portland Patrol Officers Who Were Tested and Not Tested*

| | All Officers | | | Officers With Ratings Only | | | Officers Rated and Tested | | | t (Ratings Only |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | Mean | SD | N | Mean | SD | N | Mean | SD | Vs. Rated & Tested) |
| ***Minneapolis*** | | | | | | | | | | |
| P-CRIME | 237 | 6.28 | 1.23 | 56 | 6.49 | 1.46 | 181 | 6.22 | 1.15 | 1.45 |
| P-FORCE | 237 | 6.10 | 1.19 | 56 | 6.40 | 1.35 | 181 | 6.01 | 1.13 | 2.13* |
| P-TEAMWORK | 237 | 7.54 | .95 | 56 | 7.40 | 1.35 | 181 | 7.59 | .79 | 1.27 |
| P-OVERALL | 237 | 6.94 | 1.05 | 56 | 7.01 | 1.37 | 181 | 6.91 | .93 | .64 |
| ***Portland*** | | | | | | | | | | |
| P-CRIME | 108 | 5.78 | 1.46 | 58 | 5.62 | 1.55 | 50 | 5.97 | 1.35 | 1.24 |
| P-FORCE | 109 | 5.95 | 1.20 | 60 | 5.87 | 1.11 | 49 | 6.05 | 1.32 | .76 |
| P-TEAMWORK | 108 | 6.31 | 1.61 | 60 | 6.03 | 1.74 | 48 | 6.65 | 1.36 | 2.02* |
| P-OVERALL | 109 | 5.87 | 1.42 | 60 | 5.64 | 1.38 | 49 | 6.16 | 1.42 | 1.93* |

*Significant at $p < .06$.

TABLE 47

*Means and Standard Deviations of Ratings on Overall Criterion Dimensions According to City and Police Functions (White Males Only)*

| | Patrol Officers | | | Detectives | | | Sergeants | | | Middle Command | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Mean | SD | N | Mean | SD | N | Mean | SD | N | Mean | SD |
| Cincinnati | 105 | 6.10 | 1.42 | 43 | 6.50 | 1.20 | 65 | 6.94 | 1.19 | 24 | 7.09 | 1.24 |
| Des Moines | 95 | 6.10 | 1.17 | .. | .... | .... | 38 | 5.88 | 1.55 | 20 | 5.71 | 1.14 |
| Miami | 58 | 6.41 | 1.42 | .. | .... | .... | 45 | 6.28 | 1.31 | 30 | 6.80 | .95 |
| Minneapolis | 177 | 6.90 | .95 | 45 | 6.24 | 1.33 | 38 | 6.63 | 1.26 | 28 | 5.83 | 1.23 |
| Portland | 58 | 6.15 | 1.45 | 46 | 7.12 | .99 | .. | .... | .... | 18 | 6.61 | 1.10 |
| San Diego | ... | .... | .... | 52 | 6.94 | .86 | 44 | 6.84 | 1.35 | 30 | 6.68 | 1.00 |
| Washington, D. C. | 64 | 6.32 | 1.22 | 46 | 6.61 | 1.08 | 60 | 6.22 | 1.65 | 44 | 7.12 | 1.00 |
| Combined Cities | ... | .... | .... | 55* | 6.48* | 1.18* | .. | .... | .... | .. | .... | .... |

*These values are for detectives from Albuquerque, Tucson, and Des Moines combined.

TABLE 48

*Numbers of Men, Women, Whites and Non-Whites According to Police Function*

| | Patrol Officers | Detectives | Sergeants | Middle Command |
|---|---|---|---|---|
| Men | 747 | 424 | 373 | 214 |
| Women | 13 | 26 | 9 | 0 |
| White | 640 | 377 | 334 | 208 |
| Non-White | 123 | 73 | 44 | 6 |

TABLE 49

*Means and Standard Deviations on Criterion Dimensions for White and Non-White Patrol Officers*

| Criterion Dimension | White N* | White Mean | White SD | Non-White N* | Non-White Mean | Non-White SD |
|---|---|---|---|---|---|---|
| P-CRIME | 496 | 6.19 | 1.28 | 84 | 5.60 | 1.67 |
| P-FORCE | 500 | 6.01 | 1.25 | 88 | 6.27 | 1.25 |
| P-TEAMWORK | 404 | 7.17 | 1.25 | 88 | 6.40 | 1.55 |
| P-OVERALL | 557 | 6.42 | 1.27 | 103 | 5.64 | 1.63 |

*Ns are smaller than those shown in Table 48 because rating information for certain cities and certain scales was discarded due to low reliabilities.

TABLE 50

*Means and Standard Deviations on Criterion Dimensions for Male, Female, White and Non-White Detectives*

| Criterion Dimension | White, Male N* | White, Male Mean | White, Male SD | Non-White N | Non-White Mean | Non-White SD | Female N* | Female Mean | Female SD |
|---|---|---|---|---|---|---|---|---|---|
| D-INVESTIGATE | 294 | 6.57 | 1.53 | 66 | 6.23 | 1.55 | 24 | 6.22 | 1.35 |
| D-INTEGRITY | 164 | 7.58 | 1.50 | 43 | 7.80 | 1.31 | 14 | 7.79 | .90 |
| D-REPORTING | 212 | 6.56 | 1.46 | 9 | 6.94 | 1.30 | 10 | 6.68 | 1.32 |
| D-OVERALL | 330 | 6.62 | 1.41 | 67 | 6.65 | 1.04 | 25 | 6.44 | .81 |

*Ns are smaller than those shown in Table 48 because rating information for certain cities and certain scales was discarded due to low reliabilities.

TABLE 51

*Means and Standard Deviations on Criterion Dimensions for White and Non-White Sergeants*

| Criterion Dimension | White N* | White Mean | White SD | Non-White N* | Non-White Mean | Non-White SD |
|---|---|---|---|---|---|---|
| S-TRAINING | 314 | 6.39 | 1.64 | 42 | 5.80 | 1.92 |
| S-SUPERVISION | 309 | 6.88 | 1.55 | 43 | 6.72 | 1.81 |
| S-INSPECTION | 238 | 6.48 | 1.74 | 40 | 6.24 | 1.89 |
| S-OVERALL | 312 | 6.53 | 1.42 | 43 | 6.00 | 1.61 |

*Ns are smaller than those shown in Table 48 because rating information for certain cities and certain scales was discarded due to low reliabilities.

TABLE 52

*Means and Standard Deviations on Criterion Dimensions for White Middle Level Command Officers*

| Criterion Dimension | N* | Mean | SD |
|---|---|---|---|
| C-ADMINISTRATION | 204 | 6.62 | 1.50 |
| C-COMMUNICATIONS | 186 | 6.62 | 1.55 |
| C-PERSONNEL | 186 | 6.83 | 1.62 |
| C-DEDICATION | 157 | 6.79 | 1.45 |
| C-OVERALL | 206 | 6.64 | 1.21 |

*Ns are smaller than those shown in Table 48 because rating information for certain cities and certain scales was discarded due to low reliabilities.

white sergeants on all performance scales. The greatest difference relative to the standard deviations occurred on the OVERALL scale, where the mean difference amounted to .35 SD, corresponding to an overlap coefficient of 86%. [See Dunnette (1966), pages 142–144, for an explanation of the overlap coefficient.]. Largest differences between the means for white and non-white patrol officers occur on the TEAMWORK and OVERALL scales, where the differences correspond respectively to overlap coefficients of 78 percent and 79 percent. Mean differences between whites and non-whites and between males and females are minor and of little

consequence on the criterion dimension scales for detectives. Differences shown between whites and non-whites for the patrol officer and sergeant ratings are discussed more fully later in this section where we discuss subgroup differences on the predictor scales.

d. *Height, weight, age, length of service.* Criterion dimension comparisons were also made between persons above and below the midpoints on the variables of height, weight, age, and total length of police service. Means and standard deviations on the Overall criterion dimensions are shown in Table 53. The only mean differences of consequence relate to the age and length of service variables. As might be expected, older, more experienced patrol officers received higher overall ratings than younger, less experienced officers—though the mean difference between the two subgroups is really not very large (about one-third of a standard deviation). Interestingly, the relatively younger officers in supervisory (sergeant) and command positions received higher overall ratings than relatively older officers. Apparently many of the "old hands" are also seen as "over the hill." The difference between old and young, long experience and briefer experience is particularly pronounced among command level officers.

2. *Validities by subgroup.*

a. *Cognitive tests.* Correlations were computed between scores on each of the four cognitive measures and all criterion dimension ratings, including the four pseudo-criterion RANDOM distributions. Correlations were computed separately for blacks and whites for the patrol officer, detective, and sergeant functions. The ranges of coefficients obtained for each function and median values are shown in Table 54. At first glance, none of the tests seem to hold much promise. However, a more careful look suggests that certain of the tests may merit further study. For example, the verbal and symbolic reasoning tests yield potentially useful validities for blacks in detective jobs. This is shown in the following:

| | INVESTIGATE (N=48) | INTEGRITY (N=35) | OVERALL (N=49) |
|---|---|---|---|
| Verbal Reasoning | .00 | .01 | .23 |
| Symbolic Reasoning | .30 | .31 | .37 |

The symbolic reasoning test appears especially good. As will be seen shortly, the predictor scales of the detective CAREER INDEX do not appear to show good validities for blacks. Thus, the symbolic reasoning test definitely merits further study.

Additionally, the visual speed and verbal reasoning tests appear potentially useful for black sergeant and for middle command officers, as follows:

Unfortunately, both tests correlate near zero with the overall rating for black sergeants. However, the correla-

*Black Sergeants*

| | TRAINING (N=33) | SUPERVISION (N=34) | INSPECTION (N=31) | OVERALL (N=34) |
|---|---|---|---|---|
| Visual Speed | .24 | −.02 | .33 | .03 |
| Verbal Reasoning | −.05 | .21 | .20 | −.06 |

*Middle Command Officers*

| | ADMIN (N=204) | COMM (N=186) | PERS (N=186) | DEDIC (N=157) | OVERALL (N=205) |
|---|---|---|---|---|---|
| Visual Speed | .11 | .15 | .11 | .15 | .22 |
| Verbal Reasoning | .16 | .22 | .17 | .11 | .22 |

tions against the INSPECTION ratings show some promise.

Also note the generally positive correlations between the cognitive measures and various of the predictor scales of the CAREER INDEX batteries. For the patrol officer predictor scales, highest correlations are shown with the verbal comprehension and verbal reasoning tests; of the two detective scales, DI correlates .19 with verbal comprehension; neither of the sergeant scales correlates highly with any of the tests, but the visual speed and verbal reasoning tests show uniformly positive correlations with the command level predictor scales.

b. *Validity by city.* Correlations between predictor scales and each of the criterion dimension ratings (including the pseudo-criterion RANDOM) were computed separately for officers for each of the participating cities. Tables showing these results are given in Appendix H. Results shown there are for white officers only. A later comparison shows validities separately for whites and for racial and sexual minority groups.

The validities across cities are remarkably similar to the validities for the overall sample. Not only are the actual magnitudes of coefficients similar from city to city, but the patterns of higher and lower values are very similar over all the cities.

For patrol officers, Miami is the only city with some-

TABLE 53

*Means and Standard Deviations on Overall Performance Ratings for Officers Above and Below the "Midpoint"* on Height, Weight, Age, and Length of Police Service (White Males Only)*

| | Patrol Officers | | | Detectives | | | Sergeants | | | Middle Command | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Height* | N | Mean | SD | N | Mean | SD | N | Mean | SD | N | Mean | SD |
| Shorter | 198 | 6.46 | 1.28 | 116 | 6.62 | .99 | 83 | 6.32 | 1.38 | 46 | 6.61 | 1.23 |
| Taller | 359 | 6.39 | 1.27 | 213 | 6.62 | 1.21 | 228 | 6.62 | 1.42 | 151 | 6.64 | 1.19 |
| *Weight* | | | | | | | | | | | | |
| Heavier | 291 | 6.45 | 1.28 | 195 | 6.56 | 1.20 | 190 | 6.65 | 1.38 | 126 | 6.66 | 1.20 |
| Slimmer | 266 | 6.38 | 1.27 | 133 | 6.72 | 1.05 | 120 | 6.35 | 1.47 | 71 | 6.62 | 1.20 |
| *Age* | | | | | | | | | | | | |
| Older | 266 | 6.56 | 1.20 | 156 | 6.59 | 1.25 | 161 | 6.38 | 1.46 | 99 | 6.33 | 1.26 |
| Younger | 283 | 6.30 | 1.32 | 174 | 6.66 | 1.03 | 149 | 6.68 | 1.35 | 98 | 6.94 | 1.04 |
| *Length of Service* | | | | | | | | | | | | |
| Longer | 237 | 6.63 | 1.04 | 158 | 6.52 | 1.24 | 156 | 6.45 | 1.47 | 90 | 6.32 | 1.30 |
| Shorter | 307 | 6.26 | 1.34 | 171 | 6.71 | 1.03 | 154 | 6.59 | 1.37 | 107 | 6.90 | 1.02 |

*The "midpoint" was chosen to be the category closest to the median point. Splits were done separately for each function.

what larger discrepancies from results for the total sample than other cities show. But even Miami shows acceptable validities for the PIH and PIV scales.

For detectives, the DI scale shows uniform results across all cities, but the DII scale does not hold up well for officers in Cincinnati and San Diego. Recall that the DII scale possesses lower reliability relative to the DI scale. Also, the DI scale is broader in its coverage, longer, and was shown in our analyses to be related to a wider range of performance effectiveness dimensions for detectives than the DII scale.

For sergeants, results are exceptionally uniform across all cities for both SI and SII scales and for all criterion dimensions.

For middle command officers, the validities are uniform across cities with the exception of scale CIII in Minneapolis, scale CIV in Washington, and scale CV in Des Moines. The most puzzling of these probably is scale CIII, the modal response key of the Situational Judgment Inventory. We can offer no explanation for its apparent non-validity among the 25 Minneapolis middle level command officers.

Table 55 shows mean absolute deviations between validities in the total sample and validities in each of the cities for each of the predictor scales, separately for validities against real performance ratings and against the RANDOM distributions. The mean absolute deviations based on the RANDOM distributions can be regarded as rather crude empirical estimates of the amount of variation in sizes of validity coefficients to be expected by strictly chance. Any mean deviation against actual performance ratings that exceeds by very much the amount

TABLE 54

*Range of Values and Median Values* for Correlations Between Cognitive Tests and Criterion Dimension Ratings and Between Cognitive Tests and Predictor Scales of the Career Index*

| | White Officers | | Black Officers | | RANDOM | | Predictor Scales | |
|---|---|---|---|---|---|---|---|---|
| | Range | Mdn. | Range | Mdn. | Range | Mdn. | Range | Mdn. |
| Patrol Officers (N=513→698) | −05→+09 | +01 | −24→+12 | −08 | −23→+04 | −06 | 00→+33 | +14 |
| Detectives (N=212→415) | −09→+28 | +04 | −12→+37 | +15 | −24→+25 | −03 | −09→+19 | +08 |
| Sergeants (N=282→361) | +01→+17 | +06 | −42→+33 | −04 | −13→+06 | −04 | +04→+15 | +11 |
| Middle Command (N=156→206) | −08→+22 | +11 | ..... | ... | −10→+05 | 00 | −17→+28 | +11 |

*Decimals omitted.

shown against the RANDOM scale might, therefore, be regarded as indicative of some systematic factors affecting a specific scale's validities in different cities. Using this line of reasoning leads us to conclude that predictor scales PIV, DI, CIII, CIV, and CV probably show variations in validity from city to city somewhat greater than might be expected purely by chance. Of these, CIII, CIV, and CV already have been singled out as showing unusually high deviations from the total sample values in the cities of Minneapolis, Washington, and Des Moines, respectively. If the information for those three cities is deleted for those particular scales and mean absolution deviations are recomputed, the values become 16, 10, and 14 for CIII, CIV, and CV, respectively. These values are very close to the values of 16, 10, and 11, respectively, for the CIII, CIV, and CV scales against the RANDOM criterion.

Our net conclusion from these city-by-city comparisons of the validities of the predictor scales is that they are impressively consistent. The relative stability of these values is especially reassuring in view of the rather large mean differences between various cities on ratings of job performance, as shown in Table 47. Those mean differences undoubtedly resulted in selecting items for predictor scales that might be at least partially confounded by departmental designation or geographic location. Conceivably, the predictor scales could have shown excellent validities in the total samples but poor validities in most cities because of such factors as restriction in range, systematic departmental differences in the way officers responded to the inventories, or any number of other systematic but irrelevant differences from city to city. Fortunately, these problems did not occur. It is reassuring indeed to note the relative uniformity of the predictor scale validities across all cities. To a degree, these findings of reasonably consistent validities across cities confirm the wisdom of our earlier decision that criterion mean differences probably reflected actual differences in effectiveness from city to city instead of being any artifact of the way the rating scales were used.

TABLE 55

*Mean Absolute Deviations* Between Total Sample Validities and Validities for Separate Cities According to Predictor Scale*

| Predictor Scale | Against Actual Performance Rating Scales | Against Pseudo-Criterion: Random Distribution |
|---|---|---|
| PI | 09 | 12 |
| PII | 12 | 13 |
| PIII | 10 | 09 |
| PIV | 14 | 07 |
| DI | 12 | 06 |
| DII | 16 | 17 |
| SI | 09 | 13 |
| SII | 08 | 17 |
| CI | 14 | 20 |
| CII | 11 | 15 |
| CIII | 20 | 16 |
| CIV | 14 | 10 |
| CV | 18 | 11 |

*Decimals omitted.

c. *Validity by race and by sex.* Correlations between predictor scales and each of the criterion dimension ratings (including the pseudo-criterion Random) were computed separately for officers of racial minorities (non-whites) and for women. Sufficient numbers of racial minorities were available for only the functions of patrol officer, detective, and sergeant. A sufficient number of female officers were available only for the detective function. Results of these computations are shown in Tables 56, 57, and 58. Also shown in the tables are separate validities for whites and non-whites in Washington, D.C. Washington was the only city where sufficiently large numbers of minority personnel were tested to allow a direct comparison between whites and non-whites.

The validities for the patrol officer predictor scales, shown in Table 56, are very nearly identical for non-white officers and for the officers in the total sample. Overall scales and criterion dimensions (excluding RANDOM), the median validity in the total sample is .29. The corresponding value for non-white officers also is .29. Pair-by-pair comparisons between total sample and non-white validities show that only nine of sixteen differences are greater than an absolute value of .05. Of these, validities are five times higher for non-whites—four times higher than for the total sample. Direct comparisons between validities for white and non-white patrol officers in Washington, D.C. yield similar conclusions. Rather large differences occur for only the PIII and PIV scales; for both scales, the higher validities are obtained for non-whites.

The validities for the detective predictor scales, shown in Table 57, are uniformly lower for minority officers than for officers in the total sample. The average algebraic deviation between the two sets of validities is −.11. The same result is apparent in the comparison of validities for white and non-white detectives in Washington, D.C.—especially for the DI scale, where the average discrepancy is a massive −.32 for the two criterion scales of INVESTIGATION and OVERALL. These results, even though based on a relatively small sample of minority officers, suggest that the detective predictor

## TABLE 56

*Validities* of Patrol Officer Predictor Scales Against All Criterion Dimensions for Total Sample and for Minority Officers*

| Predictor Scale | Criterion Dimension CRIME | FORCE | TEAMWORK | OVERALL | RANDOM |
|---|---|---|---|---|---|
| *PI* | | | | | |
| Total Sample (N=502) | 18 | 40 | 13 | 24 | 01 |
| All Minority (N=78–97) | 20 | 36 | 14 | 29 | 14 |
| D. C.** White (N=60–62) | 20 | 46 | −09 | 32 | −04 |
| D. C. Minority (N=72–76) | 19 | 40 | 13 | 21 | 11 |
| *PII* | | | | | |
| Total Sample (N=452) | 29 | 14 | 32 | 37 | 03 |
| All Minority (N=76–94) | 27 | 01 | 26 | 38 | 13 |
| D. C. White (N=60–62) | 36 | 32 | 28 | 49 | −16 |
| D. C. Minority (N=70–74) | 27 | 05 | 27 | 33 | 12 |
| *PIII* | | | | | |
| Total Sample (N=528) | 41 | 11 | 28 | 42 | 04 |
| All Minority (N=59–71) | 38 | 03 | 28 | 48 | −01 |
| D. C. White (N=55–57) | 23 | 18 | 02 | 28 | −07 |
| D. C. Minority (N=53–54) | 35 | 06 | 25 | 50 | 00 |
| *PIV* | | | | | |
| Total Sample (N=372) | 37 | 06 | 54 | 48 | 03 |
| All Minority (N=59–71) | 44 | 19 | 48 | 57 | 06 |
| D. C. White (N=56–58) | 03 | −19 | 51 | 24 | −06 |
| D. C. Minority (N=53–54). | 44 | 24 | 45 | 57 | 05 |

*Decimals omitted.
**D. C. refers to Washington, D. C.

## TABLE 57

*Validities* of Detective Predictor Scales Against All Criterion Dimensions for Total Sample, Minority and Female Officers*

| Predictor Scale | Criterion Dimension** INVESTIGATION | INTEGRITY | OVERALL | RANDOM |
|---|---|---|---|---|
| *DI* | | | | |
| Total Sample (N=245) | 52 | 23 | 59 | 00 |
| All Minority (N=39–63) | 33 | 15 | 42 | −06 |
| All Female (N=13–24) | 20 | 35 | 41 | −01 |
| D. C.*** White (N=36–42) | 37 | 06 | 64 | 12 |
| D. C. Minority (N=36) | 15 | 09 | 23 | −10 |
| *DII* | | | | |
| Total Sample (N=263) | 11 | 62 | 17 | −09 |
| All Minority (N=30–48) | 04 | 43 | 21 | 04 |
| All Female (N=14–22) | 04 | 60 | 16 | −33 |
| D. C. White (N=39–45) | 31 | 49 | 19 | −13 |
| D. C. Minority (N=27) | 12 | 45 | 32 | 11 |

*Decimals omitted.
**Results are not shown for the criterion dimension "Reporting" because sample sizes were only nine for minority and seven for females.
***D. C. refers to Washington, D. C.

TABLE 58

*Validities* of Sergeant Predictor Scales Against All Criterion Dimensions for Total Sample and for Minority Officers*

| Predictor Scale | Criterion Dimension TRAINING | SUPERVISION | INSPECTION | OVERALL | RANDOM |
|---|---|---|---|---|---|
| *SI* | | | | | |
| Total Sample (N=271) | 45 | 55 | 48 | 59 | 00 |
| All Minority (N=37–40) | 44 | 65 | 56 | 61 | 06 |
| D. C. White** (N=55) | 38 | 50 | 37 | 62 | 21 |
| D. C. Minority (N=23–24) | 44 | 67 | 55 | 61 | –02 |
| *SII* | | | | | |
| Total Sample (N=286) | 54 | 37 | 40 | 54 | 06 |
| All Minority (N=37–40) | 55 | 23 | 24 | 50 | –03 |
| D. C. White (N=57) | 58 | 43 | 37 | 62 | 30 |
| D. C. Minority (N=23–24) | 72 | 29 | 33 | 57 | 04 |

*Decimals omitted

**D. C. refers to Washington, D. C.

scales may be inappropriate for use with minority candidates. Validities of the DI scale also are lower for females, but the validities of the DII scale show the same pattern of results against the criterion dimensions for females and for all officers in the total sample.

The validities for the sergeant predictor scales, shown in Table 58, are very nearly identical for non-white officers and for officers in the total sample. The median validity for the former group is .53; the corresponding value is .51 for the latter group. In a pair-by-pair comparison of validity coefficients, the average absolute deviation is only .07. Direct comparisons of validities for white and non-white sergeants in Washington, D.C. yield equally impressive results. Actually, in the Washington samples, the validities are slightly higher for non-whites than for whites, showing a mean algebraic difference of +.04.

These results are extremely gratifying. Validities are essentially the same for whites and for non-whites for the predictor scales of the patrol officer CAREER INDEX and the sergeant CAREER INDEX. In contrast, the applicability of the detective predictor scales to minority and female candidates is somewhat questionable. They can perhaps be used but interpreted with caution until more information concerning their validities for minority and female personnel becomes available.

3. *Predictor Scale Scores by Subgroup.*

a. *Unfair discrimination by race.* Even though validities for tests may be the same or highly similar for minority and non-minority persons (as has been shown for the patrol officer and sergeant predictor scales), use of those tests for employment or promotion decisions could still lead to charges of unfair discrimination on the basis of race. This could be the case if minority personnel systematically score lower than non-minority personnel on the tests. Under those conditions, a higher proportion of non-minority candidates than of minority candidates would be accepted for position openings. Such disparity of selection proportions may not, in and of itself, be evidence of unfair discrimination if the tests are indeed equally valid for both minority and non-minority candidates, and it is possible to show that the lower scores obtained by minority candidates are, in fact, mirrored by similarly lower levels of job performance for minority persons. In contrast, if lower test scores for minorities are not mirrored by lower job performance ratings, a charge of unfair discrimination on the basis of test information is likely to be quite viable. In essence, then, a test developer needs to: (1) evaluate the relative levels of job performance for minority and non-minority persons; (2) compute test validities separately for minority and non-minority persons; and (3) evaluate the relative levels of test performance for minority and non-minority persons. If the validities of tests are shown to be the same for both groups, use of the test for selection or promotion decisions can be regarded as fair if the relative levels of test performance for minority and non-minority persons is equivalent to the relative levels of job performance shown by the two subgroups. Then we can legitimately state that persons, regardless of subgroup membership, with equal probabilities of showing particular levels of job performance effectiveness, will also have equal likelihoods of being "accepted" by the test results.

We have already shown that the validities of CAREER INDEX predictor scales for patrol officer and sergeant functions are very similar for minority and non-minority persons. Recall also that we showed (Tables 49, 50, and 51) that minority patrol officers and minority sergeants received lower mean performance effectiveness ratings than their non-minority counterparts in these two func-

TABLE 59

*Means* and Standard Deviations on Patrol Officer Career Index Predictor Scales According to Race of Examinee*

| Predictor Scale | Non-Minority Officers N | Mean | SD | Minority Officers N | Mean | SD | Mean Difference SD Ave. |
|---|---|---|---|---|---|---|---|
| PI | 500 | −3.73 | 8.04 | 97 | −5.01 | 6.83 | 0.17 |
| PII | 493 | −25.37 | 19.53 | 94 | −42.63 | 22.03 | 0.83 |
| PIII | 476 | −6.74 | 7.48 | 71 | −11.99 | 7.90 | 0.68 |
| PIV | 464 | −7.37 | 10.71 | 71 | −26.92 | 12.47 | 1.68 |

*Scores are in terms of raw scores based directly on the weights assigned to item responses.

TABLE 60

*Means* and Standard Deviations on Detective Career Index Predictor Scales According to Race of Examinee*

| Predictor Scale | Non-Minority Officers N | Mean | SD | Minority Officers N | Mean | SD | Mean Difference SD Ave. |
|---|---|---|---|---|---|---|---|
| DI | 312 | .32 | 16.60 | 63 | −5.22 | 14.05 | 0.54 |
| DII | 275 | −1.75 | 6.08 | 48 | .81 | 6.16 | 0.42 |

*Scores are in terms of raw scores based directly on the weights assigned to item responses.

TABLE 61

*Means* and Standard Deviations on Sergeant Career Index Predictor Scales According to Race of Examinee*

| Predictor Scale | Non-Minority Officers N | Mean | SD | Minority Officers N | Mean | SD | Mean Difference SD Ave. |
|---|---|---|---|---|---|---|---|
| SI | 302 | 3.33 | 7.92 | 40 | 1.67 | 8.95 | 0.20 |
| SII | 301 | −1.54 | 6.13 | 40 | −4.03 | 6.66 | 0.39 |

*Scores are in terms of raw scores based directly on the weights assigned to item responses.

tions. In contrast, minority detectives received mean performance ratings almost identical to the mean ratings received by non-minority detectives. It is time, then, for us to examine the predictor scale score distributions separately for minority and non-minority persons to see how closely they may mirror the distributions of rated job performance shown in Tables 49, 50, and 51.

b. *Predictor scale mean differences by race.* Means and standard deviations of predictor scale scores are shown separately for minority and non-minority officers in Tables 59, 60, and 61.

Results for patrol officers, shown in Table 59, show that mean differences are uniformly in favor of non-minority officers, ranging from an extremely small difference on PI to a whopping difference (overlap equal to only 40%) on PIV. In contrast, the mean difference on the OVERALL criterion dimension performance rating was only about one-half of a standard deviation $\frac{\text{Mean Difference}}{\text{SD ave}} = .54$. Thus, for purposes of reporting patrol officer CAREER INDEX predictor scale scores, adjustments have been made to assure overall score distributions for minority and non-minority candidates with mean differences equivalent to that shown by the overall job performance rating.

Results for detectives, shown in Table 60, show that mean differences are again in favor of non-minority officers. Both scales show mean differences of about one-half a standard deviation. In contrast, the mean difference on the OVERALL criterion dimension performance rating was negligible $\frac{\text{Mean Difference}}{\text{SD ave}} = .02$. Thus, for purposes of reporting detective CAREER INDEX predictor scale scores, adjustments have been made to assure overall score distributions for minority and non-minority candidates with no mean differences.

Results for sergeants, shown in Table 61, show small mean differences, again in favor of non-minority officers. The differences on the two predictor scales are of the same order of magnitude as the mean differences (shown in Table 51) between non-minority and minority persons on the four criterion dimension rating scales for sergeants.

[Values of $\frac{\text{Mean Difference}}{\text{SD ave}}$ for the sergeant ratings

ranged from 0.10 and 0.13 on SUPERVISION and INSPECTION to 0.33 and 0.35 on TRAINING and OVERALL.] Thus, the distributions of predictor scale scores already mirror quite well the distributions of job performance ratings for minority and non-minority persons, and no adjustments need to be made to the score distributions for minority candidates for sergeant level positions.

## I. Standard Scoring and Interpretive Information for POLICE CAREER INDEX

1. *Patrol Officer CAREER INDEX.*

a. *Interpretation of scores.* Four predictor scales are now available for estimating a candidate's potential job performance as a patrol officer. Our research has shown that each of the four scales is moderately related to ratings of overall job effectiveness for patrol officers. Yet each of the four also is somewhat more highly related to one of the criterion dimensions than to the others. For example, scale PI correlates most highly with performance ratings involving settling disputes between citizens, using force appropriately, dealing constructively with the public, and keeping cool under pressure. Scale PII is less highly correlated with such performance areas as those just mentioned but does show higher relationships with all other facets of a patrol officer's job performance. Scale PIII correlates most highly with performance ratings involving crime prevention, detection and investigation, maintaining public safety, and conscientiousness and dedication. Finally, scale PIV correlates most highly with performance ratings involving teamwork and cooperation with other officers, with other divisions and departments, and other law enforcement agencies.

Additional detailed examination of the behavioral components of the performance ratings most highly related to each of the patrol officer predictor scales, PI, PII, PIII, and PIV, yields titles and brief descriptions of what each scale is measuring as shown below:

PI [Public Contact] measures *Personal Attitudes and Opinions** related to effectiveness in those aspects of a patrol officer's job involving dealing constructively with the public, using tact, courtesy, and understanding, keeping "cool" under pressure, maintaining composure, mediating disputes between persons effectively, and developing a good police "image" in the eyes of the public.

PII [Overall] measures *Personal Attitudes and Opinions* related to overall effectiveness as a patrol officer as shown especially in such areas as team work, investigating crime, ethical conduct, accurate reporting, and maintaining a cool and even manner in handling all aspects of the job of patrol officer.

PIII [Crime Prevention] measures *Personal Preferences and Interests* related to overall effectiveness as a patrol officer. Effectiveness includes overall value to the department as shown especially in effectiveness in preventing criminal activity, being alert to suspicious situations, detecting and investigating criminal activity, maintaining public safety, and overall conscientiousness and dedication in carrying out the job of patrol officer.

PIV [Cooperation] measures *Preferences and Personal Background and Experience* related to overall effectiveness as a patrol officer. Effectiveness includes overall personal contribution to departmental effectiveness as shown especially in effective cooperation with other officers, other divisions or departments, and law enforcement agencies in effectively preventing crime and following up on criminal activities.

b. *Standard report form for patrol officer candidates.* Even though we have outlined above how each of the patrol officer predictor scales may be interpreted according to somewhat different facets of job performance, we should not lose sight of the very high correlations which were obtained between ratings on all criterion dimensions nor of the moderately high correlations between scores on each predictor scale and ratings of overall performance effectiveness. Because of these high interrelationships, we decided to provide an estimate of TOTAL POTENTIAL as a patrol officer by combining a candidate's scores on the separate scales PI, PII, PIII, and PIV. We decided to weight the scales PII and PIV somewhat more heavily than the scales PI and PIII, because the former two scales have higher reliabilities and higher estimated cross-validities than the latter two scales. Thus, the following combining weights were chosen for deriving the TOTAL POTENTIAL estimate for the patrol officer CAREER INDEX: TOTAL POTENTIAL = PI + 1.5 PII + PIII + 1.5 PIV.

Finally, recall that the separate estimates of the expected validities for the four patrol officer predictor scales are PI: .19; PII: .26; PIII: .23; and PIV: .28. If we now assume that each scale is correlated with a construct

*The italics in these interpretive descriptions refers to the kinds of items making up each scale. "Personal Attitudes and Opinions" refers to personality inventory type items. "Personal Preferences and Interests" refers to interest inventory type items. "Personal Background and Experience" refers to personal background inventory type items. "Job Judgment" (in scales for sergeants and middle command) refers to situational judgment inventory type items.

called "Overall Patrol Officer Job Performance" at least as well as these estimates, we can use them in combination with knowledge of the correlation matrix between all pairs of predictor scales to derive an estimate of the expected validity of the TOTAL POTENTIAL composite. The value we have estimated by this means is .32*

A preliminary format for reporting scores in easily interpreted form for the patrol officer CAREER INDEX is shown as Figure 3.

Note that the report format provides results separately for each predictor scale and for the weighted TOTAL POTENTIAL composite of the four scales. The scoring system also provides capability for ranking the candidates tested by any given department and reporting each candidate's relative standing within that group. The norms used for designating the letter grade and below average categories are based on the officers who took part in our studies. The category designations are equivalent to the following percentages of officers:

| | | |
|---|---|---|
| A | Top 7% | |
| B+ | Next 9% | Top 50% of officers |
| B | Next 7% | in national norm group |
| B− | Next 8% | |
| C+ | Next 19% | |

*Statisticians may well rebel in horror against this crude approach for estimating the validity of the TOTAL POTENTIAL composite, based as it is on cross-validity estimates with incompletely known statistical properties, an assumption of a hypothetical "Overall Job Performance" construct, and rather unstable estimates of the matrix of correlations between predictors. We obviously cannot say much about the above estimated value. It provides, at this stage, merely a crude "feel" for what we may hope for in using the combined predictors. New studies need to be done to pin down with more certainty what the validity of the TOTAL POTENTIAL composite may be in actual use. The same is true, of course, for the estimates made for the detective, sergeant and middle command composites.

BELOW AVERAGE: Bottom 50% of officers in national norm group

2. *Detective CAREER INDEX.*

a. *Interpretation of scores.* Two predictor scales are now available for estimating a candidate's potential job performance as a detective. Our research has shown that scores on DI are much more broadly related to effectiveness as a detective than scores on DII. Also, DI scores are highly correlated with ratings of overall job effectiveness for detectives, whereas DII scores correlate rather low with such ratings.

Descriptions of Measures

PI [Public Contact] measures *Personal Attitudes and Opinions* related to those aspects of a patrol officer's job involving dealing constructively with the public, using tact, courtesy and understanding, keeping "cool" under pressure, maintaining composure, mediating disputes between persons.

PII [Overall] measure *Personal Attitudes and Opinions* related to overall effectiveness as a patrol officer as shown especially in teamwork, investigating crime, ethical conduct, accurate reporting, and maintaining a cool and even manner.

PIII [Crime Prevention] measures *Personal Preferences and Interests* related to preventing criminal activity, being alert to suspicious situations, detecting and investigating criminal activity, maintaining public safety, and level of conscientiousness and dedication in carrying out the job.

PIV [Cooperation] measures *Preferences and Personal Background and Experience* related to effective cooperation with other officers, other

FIGURE 3. *Sample report format for results from the Patrol Officer POLICE CAREER INDEX.*

Patrol Officer

Name of Candidate: _____

Estimated Potential Effectiveness as Patrol Officer

| | *BELOW AVERAGE* | C+ | B− | B | B+ | A |
|---|---|---|---|---|---|---|
| PI Public Contact | | | | | | |
| PII Overall | | | | | | |
| PIII Crime Prevention | | | | | | |
| PIV Cooperation | | | | | | |
| TOTAL POTENTIAL | | | | | | |

This Patrol Officer candidate ranks 7 out of 163 candidates tested in your departments.

divisions or departments, and law enforcement agencies.

Detailed examination of the behavioral components of the performance ratings most highly related to each of the two detective predictor scales yields titles and brief descriptions of what each scale is measuring as shown below:

DI [Overall] measures *Personal Background and Experience* related to overall effectiveness as an investigating officer (detective). Effectiveness includes gathering and coordinating all important information at the scene of a crime; following up throughly in investigating and prosecuting a case; interrogating suspects completely and with an open mind; knowing and using proper procedures in search, seizure, and arrest; protecting physical evidence; keeping complete records; reporting all information concisely and accurately; and exerting maximum effort at all times both on and off duty.

DII [Personal Integrity] measures *Attitudes and Personal Background and Experience* related to effectiveness in those aspects of an investigating officer's job involving fairness and integrity in dealing with all cases equally, avoiding personal prejudices and bias, working effectively with juveniles, resisting opportunities to use one's badge for personal gain.

b. *Standard report form for detective candidates.* Again we should not lose sight of the high correlations which were obtained between ratings on all the criterion dimensions of detectives' job performance. Thus, even though DI and DII do provide information about somewhat separate facets of a detective candidate's expected job performance, the two scales may still be combined to form an estimate of TOTAL POTENTIAL as a detective. In combining DI and DII, we took account of the considerably higher reliability and higher estimated cross-validity shown by DI as compared with DII. Thus, the following combining weights were chosen for deriving the TOTAL POTENTIAL estimate for the detective CAREER INDEX: TOTAL POTENTIAL = 2DI + DII.

Again, recall that separate estimates were made of the expected validities for the two detective predictor scales. They are DI: .30; and DII: .21. These estimates were used to compute a crude estimate of the expected validity of the TOTAL POTENTIAL composite. The value we estimated is .36.

A preliminary format for reporting scores in easily interpreted form for the detective CAREER INDEX is shown as Figure 4.

Descriptions of Measures

DI [Overall] measures *Personal Background and Experience* related to overall effectiveness in gathering and coordinating all important information at the scene of a crime, following up thoroughly in investigating and prosecuting a case, interrogating suspects completely and with an open mind, knowing and using proper procedures in search, seizure, and arrest, protecting physical evidence, keeping complete records, reporting all information concisely and accurately, and exerting maximum effort at all times both on and off duty.

DII [Personal Integrity] measures *Attitudes and Personal Background Experience* related to fairness and integrity in dealing with all cases equally, avoiding personal prejudices and bias, working effectively with juveniles, and resisting opportunities to use one's badge for personal gain.

FIGURE 4. *Sample report format for results from the Detective POLICE CAREER INDEX.*

Investigating Officer (Detective)

Name of Candidate: _____

Estimated Potential Effectiveness as Investigating Officer

| | BELOW AVERAGE | C+ | B− | B | B+ | A |
|---|---|---|---|---|---|---|
| DI Overall | | | | | | |
| DII Personal Integrity | | | | | | |
| TOTAL POTENTIAL | | | | | | |

This Detective candidate ranks 1 out of 64 candidates tested in your department.

3. *Sergeant CAREER INDEX*

a. *Interpretation of scores.* Two predictor scales are now available for estimating a candidate's potential job performance as a sergeant. Our research has shown that scores on SI and SII are very nearly equally and highly correlated with ratings of overall job effectiveness for sergeants. However, each is also correlated most highly with slightly different facets of job performance in the job of police sergeant. SI involves a somewhat greater emphasis on planning, coordination, and decision making. SII involves a somewhat greater emphasis on dealing effectively, in a human relations sense, with subordinate officers.

Detailed examination of the behavioral components of the performance ratings most highly related to each of the two sergeant predictor scales yields titles and brief descriptions of what each scale is measuring as shown in the following:

SI [Supervisory Coordination] measures *Job Judgment* related to overall effectiveness of officers with sergeant rank. Effectiveness includes observing, evaluating, and correcting subordinates, inspecting officers and vehicles, scheduling duty time efficiently and fairly, deploying officers and vehicles to take account of crime trends, and taking initiative and showing leadership in unusual situations where firm guidelines do not exist.

SII [Supervisory Consideration] measures *Personal Background and Experience* related to overall effectiveness of officers with sergeant rank. Effectiveness includes showing personal interest and awareness of subordinates' needs, recognizing and praising good performance, training and orienting subordinates for overall improvement of performance, using courtesy and understanding in working with other persons, and showing a constructive attitude at all times.

b. *Standard report form for sergeant candidates.* In the case of these scales, it is especially important that we remember that though they measure different facets of a sergeant's job performance, they are both basically correlated with overall job performance effectiveness. Thus, they may be combined to form an estimate of TOTAL POTENTIAL as a sergeant. In combining SI and SII, no basis exists for weighting them differentially.[They are of essentially equal reliability and validity.] Thus, TOTAL POTENTIAL = SI + SII.

The separate estimates made of the expected validities for the two sergeant predictor scales are SI: .19; and SII: .22. Computing a crude estimate of the expected validity of the TOTAL POTENTIAL composite yields a value of .25.

A preliminary format for reporting scores in easily interpreted form for the sergeant CAREER INDEX is shown in Figure 5.

4. *Middle Level Command Officer CAREER INDEX.*

a. *Interpretation of scores.* Five predictor scales are now available for estimating a candidate's job performance as a middle level command officer. Our research has shown that each of the five scales is quite highly related to ratings of overall job effectiveness for middle level command personnel. Yet, detailed examination of the behavioral components of the performance ratings most highly related to each of the command level predictor scales, CI, CII, CIII, CIV, and CV, yields titles and brief descriptions of what each scale is measuring as in the following:

CI [Administration] measures *Personal Attitudes, Opinions, Background, and Experience* related to effectiveness in those aspects of a middle command officer's job involving ad-

FIGURE 5. *Sample report format for results from the Sergeant POLICE CAREER INDEX.*

Sergeant

Name of Candidate: _____

Estimated Potential Effectiveness in Rank of Sergeant

| | BELOW AVERAGE | C+ | B− | B | B+ | A |
|---|---|---|---|---|---|---|
| SI Supervisory Coordination | | | | | | |
| SII Supervisory Consideration | | | | | | |
| TOTAL POTENTIAL | | | | | | |

This candidate for promotion to rank of Sergeant ranks 9 out or 23 candidates tested in your department.

ministrative and scheduling duties, assigning manpower, preparing and reviewing reports, keeping up-to-date records, supervising equipment and station maintenance, and organizing office procedures.

CII [Overall I] measures *Personal Background and Experience* related to overall effectiveness in all aspects of a middle command officer's job. Effectiveness includes total overall contribution to departmental effectiveness through developing, training, and motivating subordinates; handling administrative and scheduling duties; accepting responsibility through dedication and commitment; dealing courteously with the public; and taking effective command in field situations.

CIII [Overall 2] measures *Job Judgment* related to overall effectiveness in all aspects of a middle command officer's job. Effectiveness includes total overall contribution to departmental effectiveness through developing, training, and motivating subordinates; handling administrative and scheduling duties; setting a good example for subordinates; dealing courteously with the public; and taking effective command in field situations.

CIV [Dedication] measures *Background and Experience* related to effectiveness in those aspects of a middle command officer's job involving dedication, integrity, and setting a good example for subordinates; accepting responsibility; resisting opportunities for using one's position for personal gain; and projecting a picture of stature and competence as a police officer.

Descriptions of Measures

SI [Supervisory Coordination] measures *Job Judgment* related to observing, evaluating, and correcting subordinates; inspecting officers and vehicles; scheduling duty time efficiently and fairly; deploying officers and vehicles to take account of crime trends; and taking initiative and showing leadership in unusual situations where firm guidelines do not exist.

SII [Supervisory Consideration] measures *Personal Background and Experience* related to showing personal interest and awareness of subordinates' needs, recognizing and praising good performance, training and orienting subordinates for overall improvement of performance, using courtesy and understanding in working with other persons, and showing a constructive attitude at all times.

CV [Personnel] measures *Job Judgment, Personal Background, and Experience* related to effectiveness in those aspects of a middle command officer's job involving evaluating subordinates; giving them feedback on their performance through discussions, commendations, and disciplinary actions; planning and implementing training as required; and giving subordinates broad opportunities for development.

b. *Standard report form for middle level command officer candidates.* Here, too, though the predictor scales may be used to evaluate slightly different facets of a candidate's potential for serving as a command officer, the performance ratings against which the scales were validated were highly correlated, and scores on each scale are highly correlated with overall effectiveness estimates. Thus, the separate scales were combined to form a TOTAL POTENTIAL estimate for middle level command candidates just as was done for the patrol officer, detective, and sergeant scales. We decided to weight scales CIII and CIV somewhat more heavily than scales CI, CII, and CV. Our reasoning for this was based on the relatively high validities shown by CIII and CIV and on the very high correlations shown between all pairs of scales CI, CII, and CV. These high correlations already assured that they would be weighted heavily; greater weights were needed for CIII and CIV in order to give them a reasonably "equal footing" in effecting variation in the TOTAL POTENTIAL composite. Thus, the following combining weights were chosen for deriving the TOTAL POTENTIAL estimate for the middle level command officer CAREER INDEX: TOTAL POTENTIAL = CI + CII + 1.5 CIII + 1.5 CIV + CV.

The separate estimates of the expected validities for the five middle level command officer predictor scales are CI: .24; CII: .34; CIII: .32; CIV: .30; and CV: .25. Again, assuming that each scale is correlated with a construct called "Overall Middle Command Officer Job Performance" at least as well as the above estimates, we can proceed to estimate an expected validity for the TOTAL POTENTIAL composite. The value we have estimated by this method is .39.

A preliminary format for reporting scores in easily interpreted form for the middle level command officer CAREER INDEX is shown in Figure 6.

5. *Operational use of POLICE CAREER INDEX.* The POLICE CAREER INDEX consists of three inventories. The first inventory, titled Biographical and Personal Information, contains 393 items. This inventory contains all items necessary for scoring predictor scales PI, PII, PIII, PIV, DI, DII, SII, CI, CII, CIV, and part of CV. The second inventory, titled Situational Judgment Inven-

FIGURE 6. *Sample report format for results from the Middle Level Command Officer POLICE CAREER INDEX.*

Middle Level Command Officer

Name of Candidate:_____

Estimated Potential Effectiveness in Middle Level Command Position

| | BELOW AVERAGE | C+ | B− | B | B+ | A |
|---|---|---|---|---|---|---|
| CI Administration | | | | | | |
| CII Overall 1 | | | | | | |
| CIII Overall 2 | | | | | | |
| CIV Dedication | | | | | | |
| CV Personnel | | | | | | |
| TOTAL POTENTIAL | | | | | | |

This candidate for Middle Level Command ranks 6 out of 42 candidates tested in your department.

tory: Sergeants, contains the 82 situational judgment items for scoring predictor scale SI. The third inventory, titled Situational Judgment Inventory: Intermediate Commanders contains the 75 situational judgment items for scoring predictor scale CIII and the situational judgment part of predictor scale CV. The inventories can be administered quickly and easily by a competent clerical person to groups of candidates numbering as large as 75 to 100. The time required for administration is no more than two hours [the average time for a candidate is about 75 minutes] for the Biographical and Personal Information Inventory and about 45 minutes for either one of the Situational Judgment Inventories. Candidates' completed inventory booklets may then be mailed to a central scoring location where responses are keypunched, automatically scored, and interpretive profiles [see Figures 3, 4, 5, 6] printed by computer and returned to the site of testing. Turnaround time between date of testing and date of receiving completed results typically should be about three to four days at most.

Descriptions of Measures

CI [Administration] measures *Personal Attitudes, Opinions, Background, and Experience* related to effectiveness in administrative and scheduling duties, assigning manpower, preparing and reviewing reports, keeping up-to-date records, and supervising equipment and station maintenance.

CII [Overall 1] measures *Personal Background and Experience* related to overall effectiveness in developing, training, and motivating subordinates, handling administrative and scheduling duties, accepting responsibility, dealing courteously with the public, and taking effective command in field situations.

CIII [Overall 2] measures *Job Judgment* related to overall effectiveness in developing, training, and motivating subordinates, handling administrative and scheduling duties, setting a good example for subordinates, dealing courteously with the public, and taking effective command in field situations.

CIV [Dedication] measures *Background and Experience* related to dedication, integrity, accepting responsibility, and resisting opportunities for using one's position for personal gain.

CV [Personnel] measures *Job Judgment, Personal Background, and Experience* related to effectiveness in evaluating subordinates, giving them feedback on their performance through discussions, commendations, and disciplinary actions, and planning and implementing training as required.

At present, we can recommend use of the CAREER INDEX testing and scoring systems for all candidates for patrol officer, sergeant, and middle command officer positions. Since the validities of predictor scales of the detective CAREER INDEX were considerably lower for minority persons than for whites, we recommend that it be used for selection decisions for only non-minority persons at present. Its administration to minority candidates for research purposes is recommended, however, in order to develop additional information about its applicability for operational use with such candidates. In the meantime, local departments may also wish to make research use of the EAS symbolic reasoning tests, which showed validities in the .30s against ratings of performance effectiveness for black detectives.

The booklets comprising the three POLICE CAREER INDEX Inventories are included in Appendix I.

# CHAPTER III. DEVELOPMENT OF POLICE ASSESSMENT CENTER EXERCISES

Besides the POLICE CAREER INDEX discussed in earlier chapters, a second major thrust in our effort to devise valid police selection and promotion tools is the development of exercises for use in police assessment centers. Most of these exercises were designed to simulate important aspects of police work. They represent as closely as possible, within the constraints of a standardized testing situation, actual job activities performed by police officers in the four functions under study; namely, patrol officers, detectives, sergeants, and intermediate commanders (lieutenants and captains). These exercises include role-play situations in which a candidate, in the role of a police officer, must deal individually with an assessor in the role of a citizen or that of another police officer. They also include group exercises in which candidates participate in discussions designed to elicit either cooperative or competitive behavior among candidates. And they include administrative paperwork simulations requiring candidates to process information and perform administrative functions similar to those performed by patrol officers, detectives, sergeants, and intermediate commanders. There are other miscellaneous simulation exercises as well. Because they simulate actual police jobs, these exercises afford candidates the opportunity to demonstrate how well they can perform activities required for effectiveness in police work.

## A. History of the Use of Assessment Centers

Simulation exercises have been the backbone of several testing programs since World War II when German and British military psychologists developed and used simulations to help select military officers. In the United States, the Office of Strategic Services (OSS) used similar procedures for selecting intelligence operatives. Candidates taking part in the OSS testing program participated in a wide range of paper-and-pencil tests, interviews, and simulations over a period of several days. The simulations were intended to reflect aspects of field intelligence work under wartime conditions and some were therefore designed to be highly stressful.

Bray and his associates used a similar assessment program as part of their long-range study of managerial talent at AT&T (Bray, Campbell, and Grant, 1974). Beginning in 1956, they assessed a large number of new managers at AT&T through the use of multiple testing procedures including interviews, paper-and-pencil tests, and simulations. The simulations, of course, were meant to represent aspects of the manager's job and included such exercises as a business game, a group discussion, and an in-basket containing the type of written administrative materials which managers must typically process.

Testing programs like those described previously came to be called "assessment centers." They are characterized by the use of multiple testing techniques; multiple observers evaluating candidates' performance in the various tests; and, what is perhaps the major innovation of these assessment programs, the use of exercises simulating important activities of the job and eliciting behavior presumed to be closely similar to behavior on the job itself.

Since the landmark study initiated by Bray and his associates at AT&T, assessment centers for managerial candidates have been developed in several American organizations, including Standard Oil Co. of Ohio; International Business Machines Corp.; General Electric Co.; Sears, Roebuck & Co.; J. C. Penney Co.; and the Internal Revenue Service. Dunnette (1971) has reviewed in detail issues surrounding the development, use, and validity of these managerial assessment centers. Recently, the use of simulation exercises in the context of an assessment center has been extended to the area of police selection and promotion. Police assessment centers have been developed for the New York City Police Department and by the International Association of Chiefs of Police as part of their Professional Police Registry and Assessment Service.

The NYCPD Assessment Center was designed to assess the potential of high level commanders in the NYCPD for effectiveness in police management (D'Arcy and Piccolino, 1973). Consultants (associated with Rohrer, Hibler, and Replogle Institute, Inc.) who worked with the NYCPD to develop this assessment center interviewed a number of police captains to determine the nature of their job activities and their relation to the rest of the police organization. Based on their study of the captain's role in the NYCPD, the consultants

established a list of dimensions for evaluating an individual's potential for effectiveness as a police manager. Twenty-seven dimensions, distributed among five major categories, were defined. These five categories and some representative dimensions appear below (from D'Arcy and Piccolino, 1973, pp. 31–33):

1. *Problem-solving dimensions.*
   - *Problem analysis.* Grasps the source, nature, and key dimensions of a problem.
   - *Judgment.* Recognizes intuitively or otherwise significant factors and comes to a sound, practical decision.
2. *Communication dimensions.*
   - *Dialogue skills.* Effectiveness of discussion and expression in person-to-person or small group interactions.
   - *Writing skills.* Expresses ideas in writing with facility.
3. *Emotional and motivational dimensions.*
   - *Reaction to pressure.* Functions in a controlled, effective manner under stress, keeps his head.
   - *Drive.* Amount of directed, sustained energy brought to bear in accomplishing objectives.
4. *Interpersonal dimensions.*
   - *Insight into others.* Ability to proceed giving due consideration to the needs and feelings of others.
   - *Leadership.* Directs the behavior of others toward the achievement of common goals by his charisma, his insights or the assertion of his will.
5. *Administrative dimensions.*
   - *Planning.* Forward thinking, anticipates situations and problems and prepares in advance to cope with them.
   - *Commitment to excellence.* Determination that task will be done well, achieves high standards.

Next, the consultants prepared a set of simulation exercises, background interview, and paper-and-pencil tests to tap the 27 dimensions of police management effectiveness. The simulations include the following:

- *In-basket exercise.* Candidates assume the role of a newly appointed precinct commander and work through a packet of memos, letters, and other administrative materials such as would be found in a police commander's in-basket.
- *Television special exercise.* A group of four to six candidates with specially assigned roles engage in a leaderless group discussion to develop a television documentary about police work.
- *Mrs. Hall's accident exercise.* Each candidate conducts an interview to learn as much as he can about the fictional case of Mrs. Hall who had an "accident."
- *Management cases exercise.* A group of six candidates participate in a leaderless group discussion to try to resolve five major problems relevant to the NYCPD.
- *Manpower planning exercise.* Candidates first work individually to prepare solutions to a problem involving manpower planning in a police organization. They subsequently present their solutions verbally to the others and then discuss the problem as a group to develop solutions and recommendations.

The five simulation exercises constitute the bulk of the Police Assessment Center for Commanders conducted by the NYCPD. In the 2½-day testing session, they provide a major portion of the total assessment information about the likelihood that a candidate will be effective in the 27 dimensions of police management effectiveness.

Researchers associated with the International Association of Chiefs of Police utilized somewhat similar procedures in developing their Police Assessment Center (Kent, Wall & Bailey, 1974). First, they specified behavioral components and conducted task analyses for the following categories of police positions: police officer, police commander, and police executive. Then, meeting with active law enforcement professionals, they prepared lists of assessment dimensions like "judgment," "decision making," and "problem analysis." And finally, to tap these assessment dimensions, IACP researchers developed several types of simulations, including group discussion exercises and in-baskets.

The general strategy for developing assessment center simulations, as illustrated by the NYCPD and IACP procedures described briefly above, includes three main steps: (a) analyzing the position which the exercises are intended to simulate, (b) developing a set of assessment dimensions based on the position analysis, and (c) developing a set of simulation exercises to tap those dimensions. To assure content validity for the assessment center process, these three steps must be carried out to meet two critically important objectives:

1. The assessment dimensions must be closely tied to job effectiveness. They must reflect attributes and behaviors that are clearly necessary for high levels of performance on the job.
2. The exercises must be closely tied to these dimensions. They must simulate job conditions and elicit behaviors that resemble actual job behaviors.

Only if both of these objectives are met can we say with any assurance that a candidate's performance during an assessment center exercise is a reasonable indicator of the likelihood that he or she will perform effectively on the job. The procedures we carried out in the present study to develop police assessment center exercises were designed to accomplish these two objectives.

## B. Development of the Assessment Dimensions

As discussed in the report of our first year's research on this project (Heckman, Groner, Dunnette, & Johnson, 1972), we began by undertaking an extensive series of interviews and workshop meetings with police officers and officials to establish a sound foundation for the project. As a result of interviews with top officials in 32 police departments located in small, medium, and large cities throughout the country, we chose to study four key police officer functions in great detail, including the job functions of patrol officer, detective, sergeant, and intermediate commander (lieutenant and captain). These interviews, coupled with intensive analyses of the Internal Affairs Unit in the Minneapolis Police Department and published information from several other departments, yielded an enumeration and classification of police failure. With this background information about police jobs and cause of police failure, we conducted an extensive series of workshops with Minneapolis police officers (over 100 in all) in the four functions mentioned above to gather information about the basic dimensions of police officer effectiveness. Since these procedures have been presented in full detail in our first year's report (Heckman et al., 1972), they are summarized only briefly here.

Participants in each workshop included incumbents and immediate superiors of the police function under discussion. Thus, patrol officers and sergeants participated in the Patrol Officer Workshop, detectives and detective supervisors in the Detective Workshop, sergeants and lieutenants in the Sergeant Workshop, and lieutenants and captains in the Intermediate Commander Workshop. These participants wrote a large number of critical incidents of police officer performance (over 400 incidents in each of the four functions) and discussed possible dimensions of job behavior underlying an officer's performance.

The incidents were then edited by our staff to make them more understandable and to cull out incidents that did not involve behavior. Once all the incidents were edited, four or five staff members read each incident, and using the dimensions suggested by the workshop participants, they each developed a set of job behavior categories that seemed to encompass all the incidents. Four or five participants (police officers) subsequently reviewed and discussed the categories we developed. Their discussion resulted in a final set of categories or dimensions.

After developing performance rating scales for the dimensions, we tried them out on a pilot basis in rating the job performance of incumbent police officers in Detroit, Dallas, and suburban communities near Minneapolis. Forty-four citizens in the Twin Cities area participated in a workshop to discuss the relevance of these dimensions for depicting their experiences with police officers. Results from these pretests and discussion groups suggest that the dimensions, with minor revisions and elaborations, are broadly applicable for describing job performance of officers in the four police functions chosen for this project.

The assessment dimensions which we used as a basis for developing simulation exercises come directly from the police performance dimensions developed by the procedures outlined above. Our assessment dimensions are thus firmly rooted in actual, on-the-job behaviors portrayed in the hundreds of critical incidents of police effectiveness developed by officers intimately familiar with the four police functions under study. They have been reviewed and discussed by police officers in police departments located in large, medium, and small cities in different parts of the country. Therefore, the assessment dimensions are closely tied to actual police officer performance and represent aspects of police behavior critically important for job effectiveness. The assessment dimensions, which are fully defined in the report of our first year's research (Heckman et al., 1972), appear in the following for the four police functions:

1. *Patrol Officer.*
   - Crime prevention
   - Investigating, detecting, and following up on criminal activity
   - Using force appropriately
   - Dealing constructively with the public
   - Handling domestic disputes
   - Traffic maintenance and control
   - Maintaining public safety and giving first aid
   - Integrity and professional ethics
   - Commitment, dedication, and conscientiousness
   - Teamwork
   - Report writing
2. *Detective*
   - Investigating the scene of a crime
   - Arrest, search, and seizure
   - Interrogating suspects
   - Investigating a case
   - Developing and utilizing informants
   - Report writing and paperwork
   - Appearing in court
   - Public relations
   - Dealing with juveniles
   - Cooperating with other officers and divisions
   - Conscientiousness and dedication
   - Integrity and honesty
3. *Sergeant*
   - Concern for subordinates

- Scheduling, coordination, deployment, and manpower allocation
- Supervision
- Performing administrative and inspection functions
- Decision making and initiative where no firm guidelines exist
- Training and planning
- Integrity, dedication, and conscientiousness
- Dealing effectively with the public and superiors

4. *Intermediate commander.*
   - Administrative duties
   - Communications
   - Scheduling
   - Training
   - Supervision
   - Commending, disciplining, and assigning efficiency ratings
   - Field command situations
   - Public and community relations
   - Dedication, integrity, setting an example

### C. Development of the Simulation Exercises

As a first step, PDI staff members spent three days observing Minneapolis police officers performing their job activities. One staff member observed Minneapolis detectives, another observed sergeants, and a third observed lieutenants and captains. Dr. Robert Flint, who worked with us as consultant to develop simulations for Patrol Officer candidates, has had extensive experience working with Minneapolis patrol officers and therefore was already familiar with their day-to-day activities. Our observations of Minneapolis officers were extremely helpful in suggesting ideas for realistic situational exercises tapping important job aspects of these police functions.

We contacted several agencies and cities that had developed their own police assessment centers to learn about their exercises, staffing, and methods for analyzing and interpreting data, including the New York City Police Department, the Kansas City Police Department, and IACP. A staff member visited and observed the New York Police Department Assessment Center for Intermediate Commanders.

After observing the day-to-day activities of Minneapolis police officers and learning more about assessment centers developed for police departments in other cities, our staff developed tentative and preliminary ideas for exercises simulating important aspects of police work reflected in the assessment dimensions for the four police functions. We developed a total of 55 preliminary ideas. They were prepared with enough detail to describe the outlines of simulations that could be more completely elaborated in a subsequent stage of our research.

To assure that when fully developed our simulations would be appropriate for police officer candidates and would be accurate and realistic abstractions of the work activities of patrol officers, detectives, sergeants, and intermediate commanders, we asked nine leading police psychologists and police officials to review our preliminary simulation ideas. Their names are listed below:

Sgt. Richard Milne
Office of Personnel and Training
Tucson Police Department

Robert B. Mills, Ph.D.
Professor of Psychology
Department of Criminal Justice
University of Cincinnati

Michael Roberts, Ph.D.
Director, Psychological Service
San Jose Police Department

Lt. Robert A. Schwartz
Planning and Research
Bureau of Police
Portland, Oregon

T. Eisenberg, Ph.D.
Police Psychologist
San Jose Police Department

Brian S. O'Leary, Ph.D.
Personnel Psychologist
Washington, D. C. Police Department

Pierce R. Brooks
Director, Department of Public Safety
Lakewood, Colorado

Martin Reiser, Ph.D.
Department Psychologist
Los Angeles Police Department

Capt. James Shaw, Ph.D.
Personnel and Training Division
King County (Washington) Dept. of Public Safety

These individuals received descriptions of our preliminary simulation ideas by mail, together with definitions for the assessment dimensions in the four police functions. They were asked to:

- Review the performance dimension definitions for each function.

- Review our ideas for simulations tapping these performance dimensions.
- Comment on how appropriate and realistic our simulation ideas are in terms of the way things actually happen in police departments. If a situational exercise or some aspect of it violates standard operating procedures, how should we change it? What kinds of background information and circumstances can we add to make the simulations more realistic? Do the simulations represent typical and important activities in the day-to-day work of police personnel? Would candidates, especially those with little or no experience in the particular function for which they are being assessed, consider the simulations fair tests of their abilities? How much time should be devoted to each simulation when used in an operational assessment center?
- Suggest additional details that will help our simulations reflect the performance dimensions more clearly. Also, suggest how we can modify the simulations so that each will reflect a wider range of performance dimensions.
- Suggest their own ideas for situational exercises tapping the performance dimensions.

Their written comments contained many helpful suggestions for improving the exercises. They offered suggestions about specific details that were not appropriate, ideas for additional simulations, and ideas for how we could enlarge the scope of some of our simulations to include a wider range of assessment dimensions.

Working from written comments of the police psychologists and police officials, PDI staff members selected the most promising seven to ten simulation ideas for each of the four police functions. These were more fully elaborated with instructions for candidates and assessors and background details to lend realism. We prepared complete instructions to be given to candidates prior to their participation in the simulations; instructions for assessors on how to administer the exercises, set up phsyical props, and observe candidates' performance during the exercises; and details regarding background information to be provided to candidates to make the simulations as realistic as possible.

Then we invited 11 police psychologists and police officials from cities across the country to a one-day conference in Minneapolis to discuss our simulations and make suggestions for improving them. The list of conferees, some of whom had already reviewed our preliminary ideas for exercises, appears below:

Captain Allan Hoehl
Advancement and Promotion Section
New York City Police Department

Robert B. Mills, Ph.D.
Professor of Psychology
Department of Criminal Justice
University of Cincinnati

Police Chief Donald Derning
Winnetka, Illinois Police Department

Inspector Myron H. Blanch
Director of Training Division
Minneapolis Police Department

David L. Gorski, Director
Department of Public Safety
Village of Golden Valley, Minnesota

Michael Roberts, Ph.D.
Director, Psychological Services
San Jose Police Department

Lt. Robert A. Schwartz
Planning and Research
Bureau of Police
Portland, Oregon

T. Eisenberg, Ph.D.
Police Psychologist
San Jose Police Department

Brian S. O'Leary, Ph.D.
Personnel Psychologist
Washington, D. C. Police Department

Pierce R. Brooks
Director, Department of Public Safety
Lakewood, Colorado

Martin Reiser, Ph.D.
Department Psychologist
Los Angeles Police Department

All our simulations were thoroughly discussed during this conference. The attendees commented on the realism of the simulations, the extent to which simulations conformed to standard operating procedures in different police agencies, how police candidates could be expected to react to the exercises, and what kinds of behavior to look for in the exercises to distinguish candidates who perform effectively on the assessment dimensions from those who perform ineffectively. Also, the attendees made many valuable comments regarding the probable level of difficulty of the exercises for candidates and how to alter simulations to make them more realistic,

technically correct, and administratively workable. Their inputs at this stage in the development of the simulations were crucial for assuring that the exercises would simulate actual, on-the-job activities of patrol officers, detectives, sergeants, and intermediate commanders.

As a result of this conference, some exercises were eliminated because they seemed inappropriate or administratively unworkable to the police psychologist and police officials attending the one-day conference. Some exercises were combined as a result of their comments. And still other exercises were conceived during the conference as a result of their suggestions for new, additional simulations. PDI staff members continued the process of fleshing out the simulations with full details and complete instructions for both assessors and candidates. At this point, we had a total of 30 completely elaborated simulations, eight for Patrol Officer candidates, eight for Detective candidates, five for Sergeant candidates, and nine for Intermediate Commander candidates.

While preparing simulations for the patrol officer function, it soon became apparent that because candidates for Patrol Officer jobs may have had little or no prior experience or training in police work, they would probably need special introductory orientation before taking part in exercises realistically simulating aspects of the patrol officer's job. Accordingly, we took steps to develop introductory training materials for Patrol Officer candidates. Captain Myron Blanch, Director of the Training Division of the Minneapolis Police Department, worked with us to prepare training materials consisting of a series of brief lecturettes on the following aspects of patrol work reflected in the patrol officer simulations: law enforcement code of ethics, dealing constructively with the public, first aid, traffic maintenance and control, use of force, crime prevention, the preliminary investigation, and report writing.

Our strategy was to precede the actual testing of Patrol Officer candidates with a brief "micro-training" session to last approximately 2 hours. During this introductory session, candidates would be acquainted with some of the principles and procedures of the patrol officer's job. They would not be trained in any comprehensive manner, but rather provided with a general awareness of what the correct, prescribed behaviors are so that their performance during the simulation exercises can be more meaningfully observed and evaluated.

Even with these introductory micro-training materials, we felt that some aspects of patrol work were so inherently difficult that they should not be simulated too realistically for Patrol Officer candidates who did not have the benefit of extensive police academy training. For example, handling domestic disputes and knowing when and how to apply appropriate degrees of force were considered areas that would be particularly unfair to stage with inexperienced, untrained candidates interacting in a simulation exercise with live actors. We decided, therefore, to use a videotaped segment showing two patrol officers handling a domestic disturbance and using varying degrees of force. Rather than have candidates actually interact with live actors simulating a domestic disturbance, we planned to have them observe the videotaped simulation, evaluate the performance of the two officers shown on the tape, and indicate what they think they would have done differently if they were the officers on the tape.

We obtained a copy of a training tape developed by the Minneapolis Police Department showing a simulation of two patrol officers (acted by two Minneapolis police officers) handling a domestic disturbance (acted by two professional actors).

We also developed a second videotaped simulation showing two police officers responding to a call from a frightened woman claiming to have heard a prowler. This tape was prepared primarily to simulate the patrol officer assessment dimensions of Crime Prevention; Investigating, Detecting, and Following Up on Criminal Activity; and Dealing with the Public. Two Minneapolis patrol officers played the parts of patrol officers responding to the call and a professional actress was hired to play the role of the frightened woman. Captain Myron Blanch and the Minneapolis Police Department provided technical assistance during the production of this videotape simulation.

The next step was to take the simulations off the drawing board and actually try them out with police officers in the four police functions. Police departments in Minneapolis, Des Moines, and Cincinnati generously agreed to cooperate with us in this stage of the development of the simulations.

We scheduled a total of 77 police officers in these three cities to participate in the pretesting of the simulation exercises. Table 62 shows the distribution of participants scheduled according to function and city. Of the 77 officers who were scheduled to participate, only approximately five or six failed to show up at the appointed times for their session. The officers participated for varying lengths of time, from 2 to 8 hours depending on what specific simulation exercises were scheduled for them.

The participating officers were either newly promoted (within the past several months) to the function in which they were scheduled to participate as pretest subjects, or they were candidates eligible for promotion to that function. Thus participants in the Intermediate Commander pretests were newly promoted lieutenants and captains and experienced sergeants eligible for promotion to lieutenant; participants in the Sergeant pretests were

TABLE 62

*Distribution of Police Officers Scheduled to Participate in the Pretesting of the Simulation Exercises, According to City and Police Function*

| City | Police Function | | | | |
|---|---|---|---|---|---|
| | Patrol Officer | Detective | Sergeant | Intermediate Commander | Totals |
| Minneapolis | 6 | 7 | 6 | 6 | 25 |
| Cincinnati | 6 | 7 | 10 | 4 | 27 |
| Des Moines | 4 | 8 | 9 | 4 | 25 |
| Totals | 16 | 22 | 25 | 14 | 77 |

newly promoted sergeants and experienced patrol officers eligible for promotion to sergeant; participants in the Detective pretests were newly promoted detectives and patrol officers eligible for promotion to detective/investigator; and participants in the Patrol Officer pretests were new patrol officers, police cadets with cadet training and several months of experience as cadets, and community service officers who also had some police training and experience but not formal police academy training or experience as full patrol officers. Several women and blacks were included among the officers participating in the pretests.

A PDI staff member began each pretest session by explaining the background and purpose of the project and enlisting the participants' help in further developing the simulations to assure that they would be realistic and fair tests for Police Officer candidates. PDI staff members administered instructions for each exercise and assumed roles that would be played by assessors in the role-play simulations.

There was no attempt to evaluate the effectiveness or ineffectiveness of participants who performed the simulations during this pretest. Rather, participants were interviewed at the end of each exercise and asked for their reactions. They were asked questions such as the following:

- Were the instructions clear? Was there any confusion about what they were to do during the exercise?
- Was there too much or not enough time allotted for the exercise?
- Did the exercise seem too difficult or too easy? Would it seem too difficult or too easy for candidates with no training or experience in the particular police function represented in the simulation? Should the amount of time allotted to the exercise or the complexity of the content of the exercise be altered to make the exercise easier or, conversely, more difficult for candidates?
- Were details of the content of the exercise realistic and technically correct? Were all the technical terms appropriate for their particular police department? Did the content of the exercise violate standard operating procedures in their department? Would inexperienced candidates understand the meaning of technical terms?
- Could they suggest any changes in the administration, instructions, or content of the exercise?
- Was there anything about the administration, instructions, or content that might be unfair to women or minority candidates?

By the end of these pretest sessions, each of the 30 simulations had been performed by an average of approximately six police officers. Their comments and reactions during and after the simulations were rich in suggestions for altering various details and improving the exercises. By and large, they seemed to feel that the simulations were generally realistic portrayals of police officers' jobs and that when used in combination the simulations would be regarded by candidates as fair tests of their abilities and potential for effectiveness.

After pooling their observations of reactions and comments made by participants in the pretests, PDI staff members once again revised the simulations. Instructions were modified for added clarity; details of the content of exercises altered for added realism; more or less time allowed for performing certain exercises for a more appropriate level of difficulty; and a few exercises dropped or replaced because they proved too unrealistic, too complex, too obvious, or too administratively awkward to be workable. These revised simulations included seven for Patrol Officer candidates, seven for Detective candidates, five for Sergeant candidates, and seven for Intermediate Commander candidates. In all, 26 simulations were finally developed. In addition, standardized background interview protocols were prepared to supplement the simulation exercises.

To summarize, PDI staff members developed 26 exercises simulating important aspects of police work in the functions of patrol officer, detective, sergeant, and in-

termediate commander by carrying out the following steps:

- Minneapolis police officers were observed as they performed their day-to-day job activities.
- Information about the exercises and administration of other police assessment centers was obtained.
- Preliminary ideas for 55 simulation exercises were prepared to tap aspects of police work reflected in the dimensions of police officer performance developed earlier during this research project.
- The simulation ideas were revised and more fully elaborated according to the written comments of nine police psychologists and police officials from a variety of police departments across the nation.
- The simulations were further revised according to the inputs of 11 police psychologists and police officials who attended a one-day conference in Minneapolis.
- Introductory micro-training materials and videotaped simulations were prepared for Patrol Officer exercises.
- The simulations were pretested with approximately 70 police officers in Minneapolis, Des Moines, and Cincinnati.
- The simulations were again revised according to the results of the pretests.

This process resulted in the final development of 26 simulation exercises for candidates seeking selection or promotion to Patrol Officer, Detective, Sergeant, or Intermediate Commander.

## D. Catalogue of Simulation Exercises

The 26 simulation exercises and related materials developed for the four police functions are summarized briefly in the following and on subsequent pages. (Full details about the content of these exercises are provided in the Assessor's Manuals which are included as appendices to this report.)

1. *Exercises and materials for patrol officer candidates.*

a. *Micro-training materials.* Candidates receive booklets containing brief lecturettes on the following aspects of the patrol officer's job: law enforcement code of ethics, dealing constructively with the public, first aid, traffic maintenance and control, use of force, crime prevention, the preliminary investigation, and report writing. Candidates are instructed to read along quietly as they listen to to the lecturettes presented on a cassette tape.

b. *Domestics.* In this exercise, candidates view a videotaped segment showing two patrol officers handling a domestic dispute. The tape is stopped at specific points and the candidates evaluate the performance of the two officers with instructions to indicate how they would have acted if they were the officers in this situation. After the tape is shown, candidates as a group discuss the officers' performance and try to achieve a consensus on the effectiveness of the performance of the officers in the videotaped segment.

c. *Frightened woman.* This is a second exercise in which candidates respond to a videotaped simulation. As in the "domestics" exercise, the tape is stopped at specific points and candidates evaluate the performance of the officers on the videotaped simulation. Then, as a group they discuss and evalute the officers.

d. *Precinct desk.* The Precinct Desk Exercise simulates two telephone calls from citizens to a police precinct. Candidates listen to a recording of these calls and hear the patrol officer at the precinct desk talk with the citizen callers. At several points, the recording is stopped and candidates are instructed to write down what they would say next if they were the officer at the desk taking the call.

e. *Traffic stop.* This is a role-play exercise in which the candidate assumes the role of a patrol officer about to issue a citation to a driver, role-played by the assessor, for failing to stop at a stop sign.

f. *Hotel call.* The Hotel Call is a role-play exercise in which the candidate assumes the role of a patrol officers dispatched to a hotel where someone who is afraid he may have hurt himself has called the police for help. The injured party is role-played by the assessor. After interviewing the victim, the candidate then writes a report of the incident.

g. *Burglary.* This is a role-play exercise in which the candidate assumes the role of a patrol officer sent to answer a call from Mr. Smith, a citizen role-played by the assessor, who called the police because his hotel room was burglarized. The candidate must go through a few steps of preliminary investigation and then write a report of this incident.

h. *Professional practices.* This exercise has been designed primarily to tap the performance dimension "Integrity and Professional Ethics." Candidates write essay answers to four questions about what would be proper for a patrol officer to do in difficult and ambiguous situations involving the exercise of professional ethics. After writing their answers for an hour, they are then instructed to discuss these four situations and decide as a group what the proper course of action in each situation would be.

i. *Background interview.* The Background Interview serves the dual purpose of (a) learning about the general background of a candidate to obtain a context for interpreting his performance in the assessment center exercises, and (b) obtaining information about his per-

formance or behavior in previous situations that might indicate his potential for effectiveness as a patrol officer. The candidate is interviewed on his present job, personal history, and aspects of his job and life experiences that pertain to his potential in specific areas of patrol officer job effectiveness.

2. *Exercises and materials for detective candidates.* The first four exercises, the "Adams Case" exercises, are a sequential series of interrelated exercises designed to simulate the investigation of a homicide. Candidates assume the role of Detective William Johnson, the detective assigned to investigate the case. During the course of these exercises, he investigates the initial scene of the crime, interviews a witness, interrogates a suspect, writes reports of his investigations and interviews, and meets with the district attorney, who, to prepare the detective for trial, asks him to answer several questions as if he were actually on the stand, first under direct questioning and then under cross-examination.

a. *Adams case: Crime scene.* The candidate is told to imagine that as Detective Johnson, he has just arrived at the scene of the crime, Marcy Adams' apartment, where he was met by two patrol officers. They told him that Marcy was found unconscious by her father, George Adams, who then called the police. The candidate has 10 minutes to investigate the scene of the crime. The candidate's task, then, is to learn as much as possible about the circumstances leading to Marcy's death by examining the bits and pieces of physical evidence in the crime scene.

b. *Adams case: Mr. Adams.* The next exercise in the Adams Case series is a role-play exercise. The candidate has 10 minutes to interview George Adams, Marcy's father, who found her unconscious in her apartment and called the police. A staff member plays the role of Mr. Adams. After examining the crime scene and interviewing Mr. Adams, the candidate writes a report of the results of his investigation. Also, he writes what his immediate next steps would be in investigating the case.

c. *Adams case: Mr. Fisher.* Next, the candidate is told that as Detective Johnson, he will interview John Fisher, a possible suspect. Fisher is role-played by an assessor. The candidate has 15 minutes to question Fisher. He then writes a report of his interview.

d. *Adams case: District Attorney.* In the final exercise in the Marcy Adams series, the candidate is instructed to imagine that four months have elapsed since he began his investigation of the Marcy Adams homicide. John Fisher has been arrested and charged with her murder; he is being tried in court in two days. To prepare Detective Johnson to testify in court, the candidate is told, the district attorney has asked to meet with him and review the case. A staff member in the role of the D.A. begins this exercise by saying that to prepare Detective Johnson for the stand, he will first ask him the kinds of questions he plans to ask under direct examination. Then he will cross-examine him as if he were the defense attorney. The candidate, in the role of Detective Johnson, is to answer these questions as if he were on the stand.

e. *Written cases exercise.* The candidate assumes the role of Daniel Boyd who has been newly promoted to the position of detective in the Homicide-Robbery-Sex Division of the Granbury Police Department. He receives files for six robbery cases, some of which have been partially investigated by another detective and others which have not yet progressed beyond the patrol officers' preliminary investigation. The candidate is told that he has been assigned these six cases. He must first establish priorities among them and, second, indicate exactly how he intends to proceed on each case.

f. *Major case planning exercise.* Candidates are divided into two groups of two (or two groups of three). Candidates in one group are instructed to assume the roles of homicide detectives from Granbury; candidates in the other group are burglary detectives from Fairfax, a nearby city. They are all instructed that a burglary occurred a few days ago in Granbury in which two patrol officers were killed. A similar burglary, unsolved, occurred in Fairfax several months earlier. Each group of candidates is given a detailed summary only of the crime which occurred in their own respective cities. The two groups are to meet as one team to plan an inter-city, cooperative investigation of these two crimes.

g. *Burglary.* The candidate assumes the role of James Hanson, a burglary detective investigating a series of household burglaries. The candidate is told that during the course of his investigations he learns that Tommy Miller, a juvenile, may be implicated as a suspect. The candidate's task is to interview Tommy's father, Mr. William Miller, role-played by the assessor. The candidate must then write a report of the interview.

h. *Background interview.* The Background Interview serves the dual purpose of (a) learning about the general background of a candidate to obtain a context for interpreting his performance in the assessment center exercises and (b) obtaining information about his performance or behavior in previous situations that might indicate his potential for effectiveness as a detective. The candidate is interviewed on his present job activities, personal history, and specific aspects of his performance in his present job that pertain to his potential for effectiveness as a detective.

3. *Exercises and materials for sergeant candidates.*

a. *In-basket exercise.* The In-basket Exercise simulates the kind of administrative paperwork a police

sergeant would have to process. The candidate assumes the role of Bill Smith, newly promoted to sergeant, replacing Sergeant Boyd who died suddenly last week. As Sergeant Smith, the candidate must take action on a number of letters, memos, and miscellaneous administrative matters that have accumulated in Sergeant Boyd's in-basket.

b. *Interview with Officer Kimple.* This is a role-play exercise in which the candidate assumes the role of Sergeant Bill Smith. His task is to meet with one of his subordinates, Officer Howard Kimple, who made an appointment with Smith through Lt. Marvin of the Personnel Department. Kimple is exploring the possibility of leaving his job for another one with a suburban police department and he wants to approach his boss, Sgt. Smith, for advice. The exercise presumes that the candidate is already somewhat familiar with information about Kimple presented in the In-basket Exercise, although he receives the relevant information again for this interview. The Kimple Interview, therefore, should follow the In-basket Exercise in the Sergeant Assessment Center.

c. *Interview with Officer Ryan.* This is another role-play exercise in which the candidate assumes the role of Sergeant Bill Smith. His task is to interview one of his subordinates, Officer John Ryan, who is role-played by the assessor. There are several indications that Ryan's performance is going downhill and that he has taken to drinking while on duty. This exercise presumes that the candidate is already somewhat familiar with information about Ryan presented in the In-basket exercise although he receives the relevant information again for his exercise. The Ryan interview, therefore, should follow the In-basket Exercise in the Sergeant Assessment Center.

d. *Supervisory style discussion.* In this exercise, candidates discuss supervisory styles of police sergeants. As a group, candidates are to agree on a particular style which is generally more appropriate for police sergeants and under what specific circumstances other styles should be used, if at all.

e. *Training exercise.* Candidates are instructed to develop porposals for a 1-hour training program or module. They are told to imagine that they will be entering their proposals in the "Five-State Police Association" competition. One proposal, from among those developed by the candidates, will be chosen as the one most deserving to be funded. Candidates can spend as much time as they wish during the evening of the first day of the assessment center (the assessment center for sergeant candidates lasts 1½ days) to prepare a course outline which will constitute their training proposal. On the next day, each candidate presents his proposal to the others in a formal, stand-up presentation. Then candidates discuss all the proposals and reach a consensus on the rank order of the training proposals according to their quality.

f. *Background interview.* The Background Interview serves the dual purpose of (a) learning about the general background of a candidate to obtain a context for interpreting his performance in the assessment center exercises and (b) obtaining information about his performance or behavior in previous situations that might indicate his potential for effectiveness as a sergeant. The candidate is interviewed on his present job activities, career aspirations, personal history, and non-work activities.

4. *Exercises and materials for intermediate commander candidates.*

a. *Background Information: Precinct Commander's files.* The Intermediate Commander Center exercises are set within the context of a precinct in the hypothetical city of Fairfax. During an introductory orientation session preceding the actual simulation exercises, candidates are given a packet of background information (Precinct Commander's files) on the hypothetical precinct and city. These materials are briefly reviewed with the candidates in the orientation session. Candidates retain the materials for reference throughout the assessment center. They are collected at the end of the last exercise.

b. *In-basket exercise.* This exercise simulates the administrative work of a precinct commander. The candidate assumes the role of Capt. Frank Roberts of the Eastside Precinct and receives a packet of materials comprising the in-basket accumulation left by Roberts' successor. Within a set period of time, the candidate must process these materials in the most effective way possible. In other words, he must make decisions, delegate, schedule, analyze reports, and communicate with subordinates, superiors, and the public.

c. *Background interview.* The Background Interview serves the dual purpose of (a) learning about the general background of a candidate to obtain a context for interpreting performance in the assessment center exercises and (b) obtaining information about performance or behavior in previous situations that might indicate the candidate's potential for effectiveness as an intermediate commander. The candidate is interviewed on present job activities, personal history, specific and pertinent aspects of present job performance, career aspirations, and non-work job activities.

d. *Lt. Pete Estes interview.* In this role-playing exercise, the candidate is asked to assume the role of Capt. Frank Roberts and conduct a performance appraisal interview with Lt. Estes. The former precinct com-

mander, Capt. Groton, has completed the annual performance rating form before his death, and the candidate must discuss these ratings with Estes according to department policy. In order to do this, the candidate is given several pieces of information in addition to the current ratings, such as Estes' personnel file and memos from the in-basket.

e. *Training presentation to recruit class.* This is a group presentation exercise in which each candidate is asked to present a 10-minute training talk to a recruit class made up of staff and other candidates. Thus, the candidate must develop a talk that will hold the "recruits" interest, while at the same time giving them useful knowledge.

f. *Sgt. John Simpson exercise.* The candidate is again asked to assume the role of Capt. Frank Roberts and must interview a formerly top-notch sergeant who has lately been doing a very poor job. Staff assessors play the role of the sergeant, John Simpson.

g. *Field command exercise.* This is an individually written exercise designed primarily to tap the performance area of field command. Candidates are again asked to assume the role of Capt. Frank Roberts of the Eastside Precinct, and are then given a packet of communications designed to test their judgment under time pressure. The candidates write down all actions they would take to each item and these actions are subsequently analyzed by staff members. For convenience, the exercise is usually given in a group setting since written communications received by the candidates must be timed with a stopwatch.

h. *Precinct reorganization exercise.* The candidate again takes on the role of Capt. Frank Roberts and must complete two assignments. The first is cued by a memo from the Chief indicating that he wishes each precinct commander to reorganize his precinct according to current crime statistics, taking into account budget restrictions, operational models used by other cities, officer morale, training, and community relations. While the candidate has some suggestions given to him, he may complete the assignment in several ways, thus allowing an assessment of his analytical skills and police knowledge, his use of all available information, and his ability to formulate a clear recommendation. Secondly, the exercise requires the candidate to conduct a staff meeting with his "subordinates" to enlist their suggestions and assistance with his reorganization.

i. *Precinct Commander group discussion.* Candidates are asked individually to formulate a plan to deal with a community relations problem and then reach a group consensus on the most effective solution. After the group discussion they (again individually) incorporate the results of the meeting into a press release for the city newspaper. Newspaper articles and incident reports are given to the candidate, along with a memo from the Chief with additional inputs.

We tried to fashion the stimulus materials associated with the exercises to be as realistic as possible. It would, of course, be possible to go even further than we did to devise additional details in some of the materials. For example, in the In-basket exercise items for sergeants and intermediate commanders and in the Written Cases items for detectives, it would be possible to make the memos, letters, reports, and other paperwork items seem more realistic by printing them on "official" forms and by including official letterheads where appropriate. Since the format of official forms, such as offense reports and letterheads, vary from one police department to another, it might be useful to have them printed specifically for the location of the testing when the assessment center simulations are used operationally.

Also, the videotapes used for some of the Patrol Officer exercises might be improved if financial resources were available to utilize more elaborate staging and higher quality technical production. The videotapes we developed and used seemed adequate, but more technically perfect tapes could be produced at a higher cost.

Tables 63 through 66 summarize the exercises developed for the four police functions by listing the exercises and showing what assessment dimensions they were designed to tap. When selecting a set of exercises for an assessment center, the exercises that seem to tap the dimensions of greatest interest should be given the strongest consideration.

## E. Pilot Tests of Police Assessment Centers

Once the simulation exercises had been individually pretested with police officers and revised accordingly, the next step was to assemble simulations into coherent packages for use in complete assessment centers of from 1 to 2 days in duration. We planned to train persons to serve as assessors and then to conduct two pilot assessment centers for each of the four police functions.

In preparation for the pilot centers, PDI staff developed videotaped recordings of all role-play simulations in the catalogue of simulations. They included the Traffic Stop, Hotel Call, and Burglary simulations for Patrol Officer candidates; the Adams Interview, Fisher Interview, District Attorney Interview, and Burglary simulations for Detective candidates; the Kimple Interview and Ryan Interview simulations for Sergeant candidates; and the Estes Interview and Simpson Interview simulations for Intermediate Commander candidates. Each role-play simulation recorded on videotape was performed by two PDI staff members who were familiar with the role requirements and who had extensive ex-

TABLE 63

*Assessment Dimensions Taped by the Patrol Officer Exercises*

| Assessment Dimensions | DOMESTICS EXERCISE | FRIGHTENED WOMAN | PRECINCT DESK | TRAFFIC STOP | HOTEL CALL | BURGLARY EXERCISE | PROFESSIONAL PRACTICES | BACKGROUND INTERVIEW |
|---|---|---|---|---|---|---|---|---|
| 1. Crime Prevention | | X | | | | | | |
| 2. Investigating, Detecting | | X | X | | X | X | | |
| 3. Using Force Appropriately | X | | | | | | | X |
| 4. Dealing with the Public | | X | X | X | X | X | | X |
| 5. Domestic Disputes | X | | | | | | | X |
| 6. Traffic Maintenance | | | | X | | | | X |
| 7. Public Safety | | | | | X | | | X |
| 8. Integrity, Ethics | | | | | | | X | X |
| 9. Commitment | | | | | | | | |
| 10. Teamwork | X | X | | | | | | |
| 11. Report Writing | X | X | | | X | X | X | |

perience in similar role-play situations for other assessment center exercises. In each case, one PDI staff member assumed the role that would be played by the candidate undergoing assessment and another staff member assumed the role that would be played by the assessor.

These videotaped role-play simulations were prepared primarily to provide a standard with which to train assessors on how to play the assessor's role in these exercises. At the same time, the tapes were intended to help train assessors on how to observe and evaluate a candidate's performance during the role-play simulations. That is, trainee assessors would be asked not only to observe the person in the role of the assessor to see how the role should be played, but also to observe the person in the role of the candidate to practice evaluating and rating candidates' performance.

PDI arranged to conduct pilot centers in four different cities so that the centers would be tested for their appropriateness in police departments serving cities that differed greatly in size, geographical location, and other characteristics. We conducted pilot centers in Chattanooga, Tennessee; Portland, Oregon; Washington, D. C.; and Minneapolis, Minnesota. Table 67 shows how the eight pilot centers were distributed among these four cities. As shown in Table 67, we conducted a total of eight pilot assessment centers—two in each police function and two in each city.

Each pilot center was staffed by three assessors under the overall guidance and administration of a PDI staff member. In each case, one assessor was a high-ranking police official at the level of captain or above in the police department hosting the center and providing candidate officers; one was a local psychologist from the city in which the center took place; and one was a local citizen working in areas like social work, religious ministry, or some other profession intimately concerned with police services from the "consumer" perspective of citizens. This mix of backgrounds among assessors was accomplished to assure that Police Officer candidates would be fairly assessed by persons with professional skills in psychology, police work, and social or community service work. A total of 18 assessors were eventually trained. Six were trained for the Intermediate Commander Centers, six for the Sergeant Centers, and six for both the Patrol Officer and Detective Centers.

Candidates who volunteered to take part in the assessment center exercises were paid for their participation. Three candidates were assessed in each pilot center. For the two Patrol Officer Centers, participating candidates were citizens who showed an interest in applying for positions as patrol officers but who as yet had no

TABLE 64

*Assessment Dimensions Taped by the Detective Exercises*

| Assessment Dimensions | CRIME SCENE & MR. ADAMS | MR. FISHER EXERCISE | DISTRICT ATTORNEY | WRITTEN CASES | MAJOR CASE PLANNING | BURGLARY EXERCISE | BACK-GROUND INTERVIEW |
|---|---|---|---|---|---|---|---|
| 1. Investigating the Scene | X | | | | | | X |
| 2. Arrest, Search, and Seizure | | X | | X | | | X |
| 3. Interrogating Suspects | | X | | X | | | X |
| 4. Investigating a Case | X | | | X | X | X | |
| 5. Developing and Using Informants | | | | | | | X |
| 6. Report Writing | X | X | | | | X | X |
| 7. Appearing in Court | | | X | | | | X |
| 8. Public Relations | X | X | | X | | | |
| 9. Dealing with Juveniles | | | | X | | X | |
| 10. Cooperating with Other Officers | X | | | X | X | | X |
| 11. Conscientiousness and Dedication | X | X | | X | | | X |
| 12. Integrity and Honesty | | | | | | | |

TABLE 65

*Assessment Dimensions Taped by the Sergeant Exercises*

| Assessment Dimensions | IN-BASKET | OFFICER KIMPLE | OFFICER RYAN | SUPERVISORY STYLE | TRAINING EXERCISE | BACKGROUND INTERVIEW |
|---|---|---|---|---|---|---|
| 1. Concern for Subordinates | X | X | X | X | X | X |
| 2. Scheduling, Coordination | X | | | | | |
| 3. Supervision | X | X | X | X | | X |
| 4. Administrative, Inspection | X | | X | | | X |
| 5. Decision Making | X | | X | X | X | X |
| 6. Training & Planning | X | | | X | X | X |
| 7. Integrity | X | X | X | X | | |
| 8. Dealing with Public & Superiors | X | | | X | X | X |

TABLE 66

*Assessment Dimensions Taped by the Intermediate Commander Exercises*

| Assessment Dimensions | IN-BASKET | BACK-GROUND INTERVIEW | ESTES INTER-VIEW | TRAINING PRESEN-TATION | SIMPSON EXERCISE | FIELD COMMAND | PRECINCT REORGANI-ZATION | PRECINCT COMMANDER DISCUSSION |
|---|---|---|---|---|---|---|---|---|
| 1. Administrative Duties | X | X | | | | | X | |
| 2. Communications | X | X | | X | | | X | |
| 3. Scheduling | X | X | | | | | X | |
| 4. Training | X | X | X | X | X | | X | X |
| 5. Supervision | X | X | | | X | X | | X |
| 6. Commending, Disciplining | X | X | X | | X | | | X |
| 7. Field Command | | X | | | | X | | |
| 8. Public Relations | X | X | | X | | | X | X |
| 9. Dedication, Integrity | X | X | X | X | | | | X |

TABLE 67

*Cities in Which the Pilot Assessment Centers Were Conducted*

- Chattanooga, Tennessee:
  Pilot Center for Intermediate Commander Candidates
  Pilot Center for Sergeant Candidates
- Portland, Oregon:
  Pilot Center for Intermediate Commander Candidates
  Pilot Center for Sergeant Candidates
- Washington, D. C.:
  Pilot Center for Patrol Officer Candidates
  Pilot Center for Detective Candidates
- Minneapolis, Minnesota:
  Pilot Center for Patrol Officer Candidates
  Pilot Center for Detective Candidates

formal training or experience. Participants in the Detective Centers were either newly promoted detectives or experienced patrol officers eligible for promotion to detective; participants in the Sergeant Centers were newly promoted sergeants or patrol officers eligible for promotion; and participants in the Intermediate Commander Centers were newly promoted lieutenants and captains or sergeants eligible for promotion. Included among the 24 candidates assessed were 17 white males, 4 black males, 1 American Indian male, and two white females.

Each pilot assessment center was conducted in three stages: (a) assessors were trained and familiarized with the procedures of the center and content of the exercises, (b) candidates performed the exercises while being observed by assessors, and (c) an integration or debriefing session was held with the assessors.

Prior to the center itself, a PDI staff member spent one day explaining the general procedures and training assessors on how to conduct exercises, observe candidates' performance, and evaluate them. This training session began with a brief introduction to the background and general concept of the assessment center technique. Then the PDI staff member reviewed the content of each exercise and described briefly how candidates might be expected to react based on the reactions of candidates participating in the pretests of the exercises.

For the role-play exercises, assessors studied the videotapes made earlier with PDI staff members in the roles to be assumed by assessors and candidates. Assessors were instructed to attend closely to the "assessor's" performance and also to take notes on the "candidate's" performance. After viewing the tapes, assessors individually rated the "candidate's" performance on the appropriate assessment dimensions. The PDI staff member then reviewed their ratings and emphasized the necessity of basing ratings and evaluations on *observed behaviors*—on what the assessor saw and heard a candidate do. Assessors practiced making behavioral observations of "candidates" performance in the videotapes and rating them accordingly.

The pilot assessment centers were held either in hotel rooms or in facilities provided by the host police department. Each center basically required a conference room

where the group discussion exercises could be held and three rooms for the role-play interview simulations.

The time required for conducting the pilot centers varied according to the function for which candidates were being assessed. Patrol Officer and Detective Centers were each scheduled to run 1 day; Sergeant Centers ran 1½ days; and Intermediate Commander Centers, 2 days. Assessors' and candidates' schedules for the pilot centers are presented in Tables 68 through 76.

TABLE 68

*Schedule for Assessors and Candidates in the Pilot Patrol Officer Center*

| Time | Activity |
|---|---|
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.—10:30 a.m. | Micro-Training |
| 10:30 a.m.—11:30 a.m. | Domestics |
| 11:30 a.m.—12:00 noon | Traffic Stop |
| 12:00 noon— 1:00 p.m. | Lunch |
| 1:00 p.m.—2:00 p.m. | Background Interview |
| 2:00 p.m.—3:00 p.m. | Professional Practices |
| 3:00 p.m.—4:00 p.m. | Frightened Woman |
| 4:00 p.m.—4:30 p.m. | Hotel Call |
| 4:30 p.m.—5:00 p.m. | Precinct Desk |

TABLE 69

*Schedule for Assessor 1 in the Pilot Detective Center*

| Time | Activity |
|---|---|
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.— 8:35 a.m. | Crime Scene (A): Instructions |
| 8:35 a.m.— 8:45 a.m. | Crime Scene (A): Examine |
| 8:45 a.m.— 8:55 a.m. | Mr. Adams (A): Interview |
| 8:55 a.m.— 9:30 a.m. | Report Writing and Rating (on A) |
| 9:30 a.m.— 9:35 a.m. | Crime Scene (B): Instructions |
| 9:35 a.m.— 9:45 a.m. | Crime Scene (B): Examine |
| 9:45 a.m.— 9:55 a.m. | Mr. Adams (B): Interview |
| 9:55 a.m.—10:30 a.m. | Report Writing and Rating (on B) |
| 10:30 a.m.—10:35 a.m. | Crime Scene (C): Instructions |
| 10:35 a.m.—10:45 a.m. | Crime Scene (C): Examine |
| 10:45 a.m.—10:55 a.m. | Mr. Adams (C): Interview |
| 10:55 a.m.—11:30 a.m. | Report Writing and Rating (on C) |
| 11:30 a.m.—12:15 p.m. | Free Time |
| 12:15 p.m.— 1:15 p.m. | Lunch |
| 1:15 p.m.— 2:15 p.m. | Free Time |
| 2:15 p.m.— 3:15 p.m. | Background Interview (C) |
| 3:15 p.m.— 3:45 p.m. | Report of Background Interview |
| 3:45 p.m.— 4:45 p.m. | Case Planning |
| 4:45 p.m.— 5:00 p.m. | Debriefing |

TABLE 70

*Schedule for Assessor II in the Pilot Detective Center*

| Time | Activity |
|---|---|
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.— 9:30 a.m. | Free Time |
| 9:30 a.m.— 9:45 a.m. | Mr. Fisher (A): Instructions |
| 9:45 a.m.—10:00 a.m. | Mr. Fisher (A): Interview |
| 10:00 a.m.—10:15 a.m. | Mr. Fisher (A): Rating on Interview |
| 10:15 a.m.—10:30 a.m. | Mr. Fisher (A): Rating on Mr. Fisher Report |
| 10:30 a.m.—10:45 a.m. | Mr. Fisher (B): Instructions |
| 10:45 a.m.—11:00 a.m. | Mr. Fisher (B): Interview |
| 11:00 a.m.—11:15 a.m. | Mr. Fisher (B): Rating on Interview |
| 11:15 a.m.—11:30 a.m. | Mr. Fisher (B): Rating on Mr. Fisher Report |
| 11:30 a.m.—11:45 a.m. | Mr. Fisher (C): Instructions |
| 11:45 a.m.—12:00 noon | Mr. Fisher (C): Interview |
| 12:00 noon—12:15 p.m. | Mr. Fisher (C): Rating on Interview |
| 12:15 p.m.—12:30 p.m. | Mr. Fisher (C): Rating on Mr. Fisher Report |
| 12:30 p.m.— 1:15 p.m. | Lunch |
| 1:15 p.m.— 3:15 p.m. | Analyze Written Cases (C) |
| 3:15 p.m.— 3:45 p.m. | Analyze Written Cases (A) |
| 3:45 p.m.— 4:45 p.m. | Case Planning |
| 4:45 p.m.— 5:00 p.m. | Debriefing |

TABLE 71

*Schedule for Assessor III in the Pilot Detective Center*

| Time | Activity |
|---|---|
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.— 9:30 a.m. | Background Interview (B) |
| 9:30 a.m.—10:15 a.m. | Report & Rating from Background Interview (B) |
| 10:15 a.m.—11:15 a.m. | Background Interview (A) |
| 11:15 a.m.—11:45 a.m. | Report & Rating from Background Interview (A) |
| 11:45 a.m.—12:15 p.m. | D.A. (with admin.) (A) |
| 12:15 p.m.— 1:15 p.m. | Lunch |
| 1:15 p.m.— 1:45 p.m. | D.A. (B) |
| 1:45 p.m.— 2:15 p.m. | D.A. (C) |
| 2:15 p.m.— 3:45 p.m. | Report Writing and Rating |
| 3:45 p.m.— 4:45 p.m. | Case Planning |
| 4:45 p.m.— 5:00 p.m. | Debriefing |

TABLE 72

*Schedule for Candidate A in the Pilot Detective Center*

| Time | Activity |
|---|---|
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.— 8:35 a.m. | Crime Scene: Instructions |
| 8:35 a.m.— 8:45 a.m. | Crime Scene: Examine |
| 8:45 a.m.— 8:55 a.m. | Mr. Adams: Interview |
| 8:55 a.m.— 9:20 a.m. | Crime Scene: Report |
| 9:20 a.m.— 9:30 a.m. | Crime Scene: Next Steps |
| 9:30 a.m.— 9:45 a.m. | Mr. Fisher: Prepare |
| 9:45 a.m.—10:00 a.m. | Mr. Fisher: Interview |
| 10:00 a.m.—10:15 a.m. | Mr. Fisher: Report |
| 10:15 a.m.—11:15 a.m. | Background Interview |
| 11:15 a.m.—11:45 a.m. | Free Time |
| 11:45 a.m.—12:15 p.m. | D.A. |
| 12:15 p.m.— 1:15 p.m. | Lunch |
| 1:15 p.m.— 3:15 p.m. | Written Cases |
| 3:15 p.m.— 3:45 p.m. | Free Time |
| 3:45 p.m.— 4:45 p.m. | Case Planning |
| 4:45 p.m.— 5:00 p.m. | Debriefing |

TABLE 73

*Schedule for Candidate B in the Pilot Detective Center*

| Time | Activity |
|---|---|
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.— 9:30 a.m. | Background Interview |
| 9:30 a.m.— 9:35 a.m. | Crime Scene: Instructions |
| 9:35 a.m.— 9:45 a.m. | Crime Scene: Examine |
| 9:45 a.m.— 9:55 a.m. | Mr. Adams: Interview |
| 9:55 a.m.—10:20 a.m. | Crime Scene: Report |
| 10:20 a.m.—10:30 a.m. | Crime Scene: Next Steps |
| 10:30 a.m.—10:45 a.m. | Mr. Fisher: Prepare |
| 10:45 a.m.—11:00 a.m. | Mr. Fisher: Interview |
| 11:00 a.m.—11:15 a.m. | Mr. Fisher: Report |
| 11:15 a.m.—12:15 p.m. | Free Time |
| 12:15 p.m.— 1:15 p.m. | Lunch |
| 1:15 p.m.— 1:45 p.m. | D.A. |
| 1:45 p.m.— 3:45 p.m. | Written Cases |
| 3:45 p.m.— 4:45 p.m. | Case Planning |
| 4:45 p.m.— 5:00 p.m. | Debriefing |

TABLE 74

*Schedule for Candidate C in the Pilot Detective Center*

| Time | Activity |
|---|---|
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.—10:30 a.m. | Written Cases |
| 10:30 a.m.—10:35 a.m. | Crime Scene: Instructions |
| 10:35 a.m.—10:45 a.m. | Crime Scene: Examine |
| 10:45 a.m.—10:55 a.m. | Mr. Adams: Interview |
| 10:55 a.m.—11:20 a.m. | Crime Scene: Report |
| 11:20 a.m.—11:30 a.m. | Crime Scene: Next Steps |
| 11:30 a.m.—11:45 a.m. | Mr. Fisher: Prepare |
| 11:45 a.m.—12:00 noon | Mr. Fisher: Interview |
| 12:00 noon—12:15 p.m. | Mr. Fisher: Report |
| 12:15 p.m.— 1:15 p.m. | Lunch |
| 1:15 p.m.— 1:45 p.m. | Free Time |
| 1:45 p.m.— 2:15 p.m. | D.A. |
| 2:15 p.m.— 3:15 p.m. | Background Interview |
| 3:15 p.m.— 3:45 p.m. | Free Time |
| 3:45 p.m.— 4:45 p.m. | Case Planning |
| 4:45 p.m.— 5:00 p.m. | Debriefing |

TABLE 75

*Schedule for Assessors and Candidates in the Sergeant Center*

| Time | Activity |
|---|---|
| DAY I | |
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.—10:00 a.m. | In-basket |
| 10:00 a.m.—10:30 a.m. | In-basket Interview |
| 10:30 a.m.—10:45 a.m. | Coffee Break |
| 10:45 a.m.—11:15 a.m. | Prepare for Ryan Interview |
| 11:15 a.m.—11:45 a.m. | Ryan Interview |
| 11:45 a.m.—12:00 noon | Ryan Report |
| 12:00 noon— 1:00 p.m. | Lunch |
| 1:00 p.m.— 3:00 p.m. | Background Interview |
| 3:00 p.m.— 5:00 p.m. | Preparation for Training Exercise |
| DAY II | |
| 8:00 a.m.— 8:15 a.m. | Prepare for Kimple Interview |
| 8:15 a.m.— 8:45 a.m. | Kimple Interview |
| 8:45 a.m.— 9:45 a.m. | Supervisory Style Discussion |
| 9:45 a.m.—10:00 a.m. | Coffee Break |
| 10:00 a.m.—10:30 a.m. | Training Presentation |
| 10:30 a.m.—11:30 a.m. | Training Discussion |
| 11:30 a.m.—12:00 noon | Debriefing |

TABLE 76

*Schedule for Assessors and Candidates in the Intermediate Commander Center*

| Time | Activity |
|---|---|
| DAY I | |
| 8:00 a.m.— 8:30 a.m. | Orientation |
| 8:30 a.m.—10:30 a.m. | In-basket |
| 10:30 a.m.—12:30 p.m. | Background Interview |
| 12:30 p.m.— 1:15 p.m. | Lunch |
| 1:15 p.m.— 1:45 p.m. | Prepare for Estes Interview |
| 1:45 p.m.— 2:15 p.m. | Estes Interview |
| 2:15 p.m.— 5:00 p.m. | Prepare for Precinct Reorganization Exercise |
| DAY II | |
| 8:00 a.m.— 8:30 a.m. | Precinct Reorganization Exercise: Staff meeting led by Candidate A with Candidates B and C as subordinates |
| 8:30 a.m.— 9:00 a.m. | Precinct Reorganization Exercise: Staff meeting led by Candidate B with Candidates A and C as subordinates |
| 9:00 a.m.— 9:30 a.m. | Precinct Reorganization Exercise: Staff meeting led by Candidate C with Candidates A and B as subordinates |
| 9:30 a.m.—10:00 a.m. | Prepare for Simpson Interview |
| 10:00 a.m.—10:30 a.m. | Simpson Interview |
| 10:30 a.m.—11:30 a.m. | Prepare for Training Presentations to Recruit Class |
| 11:30 a.m.—12:15 p.m. | Training Presentations to Recruit Class |
| 12:15 p.m.— 1:00 p.m. | Lunch |
| 1:00 p.m.— 1:45 p.m. | Field Command Exercise |
| 1:45 p.m.— 2:45 p.m. | Prepare for Precinct Commander Group Discussion |
| 2:45 p.m.— 3:45 p.m. | Precinct Commander Group Discussion |
| 3:45 p.m.— 4:15 p.m. | Write Press Release for Precinct Commander Group Discussion |
| 4:15 p.m.— 5:00 p.m. | Debriefing |

Schedules for the Patrol Officer, Sergeant, and Intermediate Commander Centers were relatively simple because all three assessors and all three candidates in each were conducting or performing the same exercises at the same time. For example, in the Sergeant Center, from 8:30 to 10:00, all three candidates completed the In-basket Exercise and from 11:00 to 11:45 Assessor I role-played Officer Ryan for one candidate while Assessor II did the same for the second candidate and Assessor III for the third candidate.

Schedules for the Detective Center, however, were more complicated. Because only one crime scene was set up (for the Adams Case exercises) and because only one candidate at a time could complete the Crime Scene Exercise, it was necessary to stagger the schedules. To facilitate scheduling, one assessor did the Mr. Adams role-play with all three candidates, another did all the Mr. Fisher role-play interviews, and the third did all the District Attorney role-play interviews.

The first item scheduled in all the pilot centers was a brief orientation session with candidates and assessors. The PDI staff member responsible for the overall administration of each center explained that the center was a pilot test being done as part of a research project funded by LEAA, introduced the assessors, described briefly the background of the assessment center technique, and explained in general terms the procedures and schedules for the center. Candidates were assured that assessors' observations and reports of their performance would be regarded as privileged and confidential information and that how they performed during the center would in no way help or hinder their careers. Also, they were told not to expect feedback from assessors on their performance.

At the end of each center, if time remained, assessors and candidates discussed their reactions to the assessment center in general and to specific exercises. Candidates provided several suggestions for improving exercise instructions and revising exercise content for increased realism.

After candidates were dismissed, assessors completed their reports and performance ratings for the various exercises. Following the Sergeant and Intermediate

Commander Centers, thorough integration sessions were conducted during which assessors discussed their observations of candidates' performance in the individual exercises and reached consensus on overall ratings. These integration sessions lasted approximately 2 hours per candidate. Finally, the PDI staff member administering the center discussed the center procedures and exercise content with assessors and elicited suggestions for improving the exercise format, administrative procedures for conducting the center, instructions provided to assessors for each exercise, rating and reporting procedures, and assessor training.

Formal integration sessions were not held following the Patrol Officer and Detective Centers. We felt that when the assessment centers become implemented and operational, so many candidates would be tested in these two lower-level centers that the cost of conducting integration sessions for Patrol Officer and Detective candidates would become prohibitive. The two upper-level centers, however, would likely have relatively fewer candidates. Moreover, the increased importance attached to correct promotional decisions for Sergeant and Intermediate Commander candidates more readily justifies the cost of thorough integration sessions following assessment centers for these two functions. After assessors completed their ratings in the Patrol Officer and Detective Centers, a PDI staff member briefly reviewed assessors' ratings and then discussed the center, exercises, and administrative procedures with assessors.

In summary, we conducted eight pilot assessment centers, two for each of the police functions under study, by carrying out the steps listed below:

- Videotape recordings of all role-play exercises were made with experienced PDI staff members in the roles of "assessor" and "candidate" to standardize the role sets for assessors and to help train assessors in observing and evaluating candidates' performance.
- Arrangements were made with police departments in four cities—Chattanooga, Tennessee; Portland, Oregon; Washington, D. C.; and Minneapolis, Minnesota—to conduct pilot centers with the cooperation of their police officers.
- Six teams of assessors were trained to conduct exercises, observe candidates, and evaluate their performance. Each team consisted of a police official at the level of captain or above, a local psychologist, and a local citizen active in some form of community or social service work. Two teams were trained for Intermediate Commander exercises, two for Sergeant exercises, and two for both Patrol Officer and Detective exercises. In all, 18 assessors were trained.
- Three "candidates" were tested in each pilot center. Candidates were either applicants for, or newly promoted incumbents of, positions in the police function represented in the assessment center exercises. In all, 24 candidates participated in the pilot centers.
- At the end of the Intermediate Commander and Sergeant Centers, integration sessions were held during which assessors pooled observations of candidates' performance and arrived at consensual judgment of candidates' effectiveness on the assessment dimensions.
- At the end of all the pilot centers, assessors were asked for their suggestions for improving the center procedures, exercise content, and assessor training procedures.

We were gratified with the general success of these pilot centers. It was encouraging that even with minimal training, assessors seemed able to make sound, behavioral observations of candidates' performance, form considered evaluations of candidates' effectiveness along the assessment dimensions backed by their behavioral observations, and arrive at consensual judgments on the relative overall levels of effectiveness shown by candidates. In addition, suggestions offered by both candidates and assessors were very helpful and to a large degree were incorporated into the final revisions of the content of the simulations and the total assessment center process.

## F. Preparation of Assessors' Manuals

Based on reactions and suggestions from persons trained to serve as assessors in the pilot assessment centers, PDI staff members prepared four comprehensive assessor's manuals, one for each of the four police functions for which we developed simulation exercises. (These assessor's manuals are included in Appendix J of this report.) The manuals contain complete definitions for the assessment dimensions tapped by the simulations; detailed instructions for assessors on how to conduct exercises, observe candidates, and evaluate their performance; sample copies of all written instructions and materials for candidates; and sample copies of all report and rating forms to be completed by assessors after each exercise. The manuals were designed to be used by persons serving as assessors when police assessment centers become operational. For convenience, pages are color-coded according to whether they are explanations and instructions for assessors, instructions and other written materials to be used by candidates, or rating forms to be completed by assessors:

- *White* pages contain explanations and instructions for assessors. For each exercise, they include a brief

summary of the exercise; a list of the performance dimensions tapped by the exercise; instructions for setting up and administering the exercise; and instructions for role-play exercises, guidelines, and information to be used by assessors in acting out their roles.

- *Blue* pages are copies of all written materials to be handed out to candidates in each exercise. They include candidates' written instructions and background information, answer and rating forms to be completed by candidates, and other written materials to be processed by candidates during the exercise.
- *Pink* pages are copies of rating forms to be completed by assessors after each exercise.

## G. Recommendations for Implementing Police Assessment Centers

Police assessment centers similar to the ones we conducted as pilot tests can be developed for any of the four police functions by selecting various combinations of exercises and simulations. Exercises should be chosen according to whether they tap assessment dimensions of greatest interest. For instance, in a center for Detective candidates, if the dimension "Interrogating Suspects" is regarded as highly important, the "Adams Case: Fisher Interview" which has been specially designed to tap that dimension should be included among the set of simulations. Similarly, if "Appearing in Court" is seen as important, the "Adams Case: District Attorney Interview" should be included. On the other hand, if "Appearing in Court" is a relatively unimportant dimension, this exercise can be excluded. The point is that assessment centers for any given application can be tailor-made to emphasize assessment dimensions of greatest concern in a particular police department by including exercises that tap those dimensions and excluding exercises that do not. Tables 63 through 66 summarize what assessment dimensions are tapped by each exercise in the catalogue of exercises and simulations and can be used as guides when selecting exercises for a tailor-made police assessment center.

The duration of an assessment center—the amount of time required to test a candidate in all exercises included in the center—will depend largely on how many exercises are included and the ratio of number of assessors to number of candidates. As a rule of thumb, assessors' time seems most efficiently utilized when an assessment center is staffed with three assessors observing a group of six candidates. This ratio of one assessor for every two candidates is prevalent in many assessment centers.

We recommend that each team of assessors include one police official at the level of captain or above in the police department wishing to select or promote candidates, one psychologist with training and experience in personnel selection and placement, and one local citizen active in some aspect of social or community service work. This combination of backgrounds and professional skills among assessors should help assure a fair and balanced assessment of police candidates.

Exercises in an assessment center should be scheduled such that candidates are observed and evaluated by different assessors in different exercises. By the end of the center, each candidate should be observed and evaluated by all three assessors, each assessor having observed a candidate in different exercises. This helps control for any bias on the part of an assessor whose evaluations of a particular candidate will be compared to those made by other assessors on the same candidate.

Persons serving as assessors will require intensive and thorough training which should include the following steps:

- Study and become thoroughly acquainted with the definitions of assessment dimensions to be measured during the assessment center.
- Review the content and procedures of all exercises included in the center as described in the Assessor's Manual.
- Observe role-play exercises as portrayed on videotaped "standards" or by experienced persons acting the roles live.
- Practice role-play exercises and become familiar with the standardized role sets to be assumed by assessors.
- Practice conducting background interviews.
- Practice recording *behavioral* observations of candidates' performance and making evaluative ratings of performance backed by behavioral observations.
- Observe experienced assessors as they conduct an actual assessment center, carry out role-play exercises and background interviews, make behavioral observations of candidates, and rate their effectiveness.

Such training for assessors will probably require 2 or 3 days plus the time required to observe an actual assessment center being conducted by experienced assessors.

After an assessment center, assessors should pool their observations of candidates' performance and discuss candidates' overall effectiveness in the assessment dimensions. At least for the Intermediate Commander and Sergeant Centers, these integration sessions should be structured to cover thoroughly all the important information learned about each candidate during the center. Assessors should discuss each other's observations and ratings for a candidate and try to reach consensus on how

effective the candidate was on each dimension. This process would require 1 or 2 hours per candidate.

For the detective and patrol officer centers, integration sessions should include at least a general discussion of candidates' relative effectiveness as shown by their assessment center performance and a final ranking of all candidates participating in the center according to their overall effectiveness. This mini-integration discussion would last a total of approximately 2 or 3 hours.

Once all assessors' reports and ratings on candidates' performances have been completed and after the integration sessions are held, a summary report on the performance of each candidate should be prepared, either by the assessor who did the Background Interview with the candidate or by the administrative coordinator of the assessment center. This summary report should contain an abbreviated version of the candidate's personal history, job background, and career aspirations as revealed in the Background Interview; a brief description of how effectively the candidate performed on each assessment dimension; and a statement of the candidate's overall effectiveness and potential in comparison to other candidates. The report should be a true representation of the observations and ratings made by assessors during the individual exercises and should reflect their consensus evaluation formed during the integration session.

# CHAPTER IV. OPERATIONAL USE OF POLICE CAREER INDEX AND REGIONAL ASSESSMENT CENTERS FOR SELECTION AND CAREER GUIDANCE OF POLICE OFFICERS

Chapters II and III have described the development and validation of POLICE CAREER INDEX inventories and the development and pilot testing of assessment center job simulations for each of four police functions. These materials are now ready and available for operational use by local departments to help them in making selection, training, promotion, job transfer, and career counseling decisions for police candidates and/or experienced police officers working in their departments.

Figure 7 shows how the POLICE CAREER INDEX Inventories can be used in concert with Regional Police Officer Assessment Centers* to provide information to local departments for use in carrying out the above personnel decision practices. The various stages shown in Figure 7 are described briefly below (The numbers on the boxes shown in Figure 7 refer to the stages discussed in the following:)

1. A local department seeking qualified candidates for vacancies would announce the availability of such positions and designate a date for administration of the POLICE CAREER INDEX (PCI) Biographical and Personal Information Inventory.
2. As candidates appear, the local department would carry out some form of preliminary screening, such as brief interviews, reference checks, etc., to develop a roster of candidates to be admitted to the PCI Inventory administration session. Sufficient numbers of PCI booklets and answer sheets would be obtained; and, on the appointed day, all qualified candidates would complete the PCI Inventory.
3. Completed answer sheets would be forwarded to a centralized computer scoring service for scoring and automatic interpretation. At that time, the local department would also indicate the types of predictor scale scores desired for each candidate. Usually a department reviewing inexperienced candidates would probably request scoring for only the patrol officer predictor scales, although early guidance and training of a newly hired trainee could perhaps profit by obtaining information relevant to the other police functions as well.

   Occasionally, a department might wish to evaluate an experienced police officer from some other department who might be under consideration for a supervisory or command job. At such times, the department would administer the appropriate PCI Situational Judgment Inventory (Sergeant or Middle Command) in addition to the PCI Biographical and Personal Information Inventory and request scoring on the appropriate keys.

   The scoring service would carry out the scoring and return interpretative profiles (similar to those shown as Figures 3, 4, 5, and 6 in Chapter II) to the local department within 2 days.
4. The PCI results would then be used in conjunction with additional departmental screening methods (such as physical examination, panel interview, etc.) to form an overall ranking of candidates.
5. The overall ranking of candidates would, for most departments, probably lead directly to the acceptance of the highest ranking persons to enter police training.

   Some departments might, however, seek further information on some of the candidates. These candidates would be asked to attend a 1-day Regional Patrol Officer or Detective Assessment Center. This option would, of course, involve considerably greater cost to the department. But in

*General procedures involved in scoring and interpreting the various predictor scales of the three inventories of the POLICE CAREER INDEX are described on pages 50–55 of Chapter II. General procedures necessary for setting up Regional Police Assessment Centers are described on pages 76–77 of Chapter III. Essentially, implementation of such regional centers requires that a cadre of persons (including police officers at level of Captain or above, psychologists, and civic minded citizens) be trained to be available to serve as staff members of the centers. Regional centers are necessary because smaller departments would have neither the resources nor the numbers of candidates necessary to warrant developing their own centers. Large departments, however, may well consider the feasibility of establishing their own centers (as detailed in Chapter 3) and making the assessment service available on a fee basis to smaller departments.

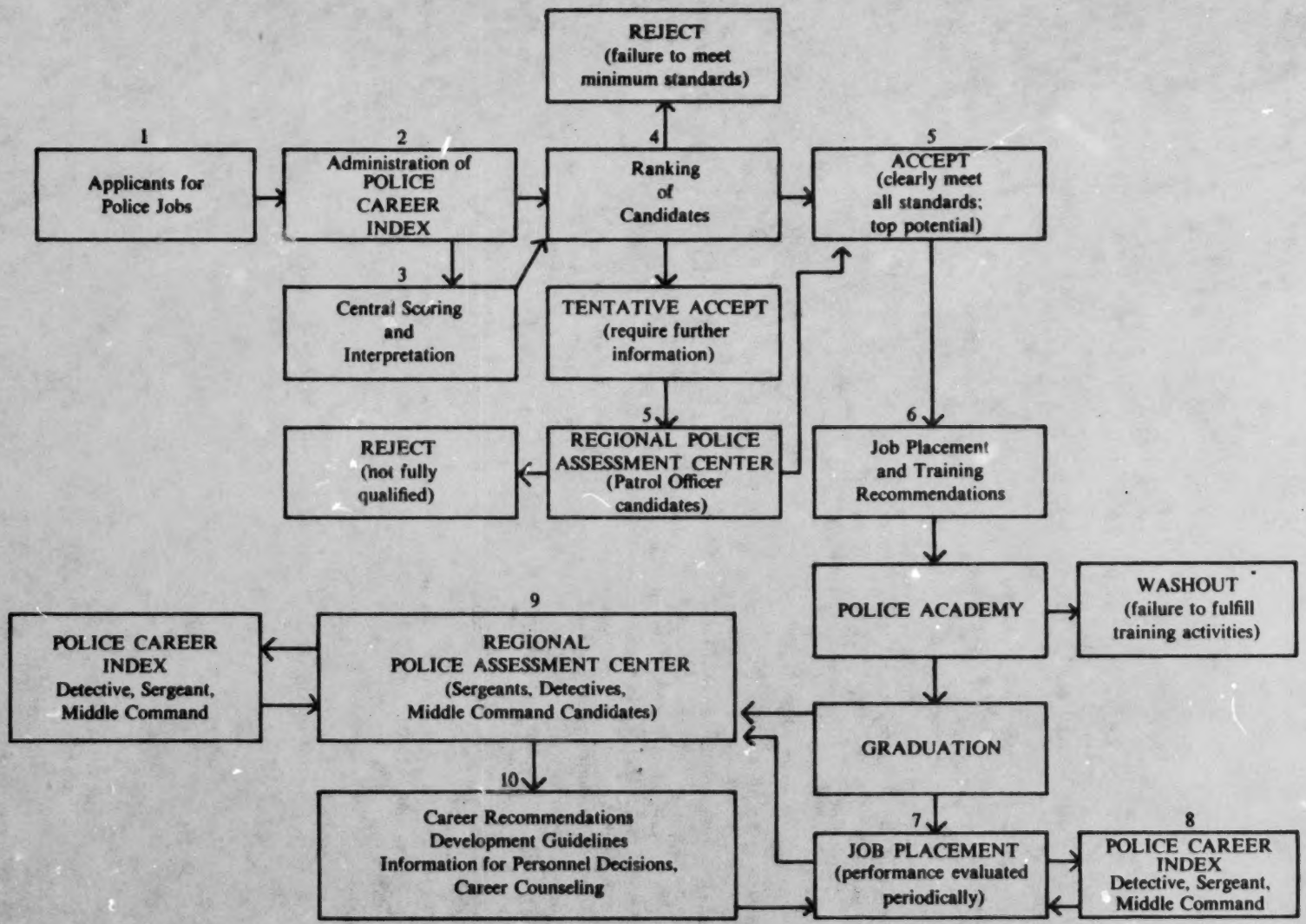



FIGURE 7. Stages of personnel selection, performance evaluation, assessment, and career placement decisions.

some instances, the richness of the behavioral information might warrant obtaining such information.

6. Each candidate finally accepted should then be interviewed in a "feedback" session where the wealth of information obtained about her or him during the selection process would be discussed, with particular emphasis placed upon implications of the information for special training needs, areas of strength and weakness, possible career opportunities, etc.
7. Upon graduation from police academy training, officers would typically be assigned directly to police jobs either as patrol officers or as detectives. Again, some departments might seek further information about an officer early in his or her career by asking him or her to attend a regional police assessment center. More typically, however, the job performance of young officers would be evaluated periodically according to the department's existing personnel practices.
8. At some point in a young officer's developing career, the department might administer various inventories of the POLICE CAREER INDEX and request scoring on predictive scales bearing on detective, sergeant, and middle command jobs. Such information then would be used in conjunction with the accumulating knowledge of the officer's job performance to build a preliminary roster of "above average potential" officers for later promotional consideration when advancement opportunities develop.
9. At some stage (or, perhaps on several occasions) in an officer's career, the department would ask that he or she attend a regional police assessment center. There, the officer's potential would be evaluated according to the simulations and behavioral observation methods described in Chapter III. If PCI information were not available on a candidate at the time of attending the regional center, the information would be obtained at that time as part of the total assessment center procedure.
10. Information about the officer's performance at the regional center would then be transmitted to appropriate persons in the department. It is desirable that the same information be given, with departmental approval, to the officer so that optimal use would be made of it in developing jointly agreed upon career recommendations, guidelines for further personal development, and basic behavioral information to be combined with all other information in developing a firm evaluation of his or her potential for serving effectively in the whole range of police positions available in the department.

The two "instruments" developed in this research program provide a total personnel evaluation system for decision making at all levels of police work, ranging from entry to the force to career guidance for individual officers and the development of increasingly accurate estimates of potential as those officers gain maturity and experience in their jobs. As has been mentioned, the central and most important feature of these two coordinated assessment methods (the POLICE CAREER INDEX and the job-focused Police Officer assessment simulation exercises) is their grounding, at every stage of development, in behaviorally explicit patterns defined by the most critical features of police officer job performance. Thus, the POLICE CAREER INDEX scales are directly interpretable according to expected patterns of police job performance instead of being in the form of the usual traits measured by psychological tests. And, the assessment simulations have been shown to elicit successfully exactly those behavior patterns discovered to be critical to successful performance in different police jobs. The total evaluation system shown in Figure 7 provides a base, therefore, for gathering and accumulating information over time that is increasingly relevant and accurate for making personnel decisions about persons and jobs in police departments.

# REFERENCES


Baehr, M. E., Furcon, J. E., & Froemel, E. C. *Psychological assessment of patrolman qualifications in relation to field performance.* Washington, D. C.: U. S. Government Printing Office, 1968.

Bownas, D. A., & Dunnette, M. D. *The development of the _____ career index.* Minneapolis: Personnel Decisions, Inc., 1975.

Bray, D. W., Campbell, R. J., & Grant, D. L. *Formative years in business.* New York: Wiley, 1974.

D'Arcy, P. F., & Piccolino, E. B. *Development of an assessment center for police commanders: New York City Police Department.* New York: Rohrer, Hibler, & Replogle Institute, Inc., 1973.

Dunnette, M. D. *Personnel selection and placement.* Belmont, Calif.: Wadsworth, 1966.

Dunnette, M. D. The assessment of managerial talent. In Paul McReynolds (Ed.) *Advances in psychological assessment: Volume Two.* Palo Alto, Calif.: Science and Behavior Books, Inc., 1971, pp. 79–108.

Ebel, R. L. Estimation of the reliability of ratings. *Psychometrika,* 1951, 16, pp. 407–424.

Eisenberg, T., & Murray, J. M. Selection. In O. G. Stahl and R. A. Staufenberger (Eds.) *Police personnel administration,* Washington, D. C.: Police Foundation, 1974.

Furcon, J., Froemel, E. C., Franczak, R. G., & Baehr, M. E. *A longitudinal study of psychological test predictors and assessments of patrolman field performance.* University of Chicago: Industrial Relations Center, 1971.

Groner, D. M., Johnson, P. D., & Dunnette, M. D. Police selection practices and research results: A review. Unpublished manuscript. Minneapolis: Personnel Decisions, Inc., 1972.

Heckman, R. W., Groner, D. M., Dunnette, M. D., & Johnson, P. D. *Development of psychiatric standards for police selection.* (Report of research under Grant Award N171–085–G) Washington, D. C.: Law Enforcement Assistance Administration, 1972.

Katzell, R. A. Symposium: The need and means of cross validation. *Educational and Psychological Measurement,* 1951, 11, pp. 1–34.

Kent, D. A., Wall, C. R., & Bailey, R. L. Assessment centers: A new approach to police personnel decisions. *The Police Chief,* June, 1974, pp. 72–77.

Landy, F. J., & Goodin, C. V. Performance appraisal. In O. G. Stahl and R. A. Staufenberger (Eds.) *Police personnel administration,* Washington, D. C.: Police Foundation, 1974.

Nie, N. H., Bent, D. H., & Hall, C. H., *Statistical package for the social sciences.* New York: McGraw Hill, 1970.

Ruch, F. L., & Ruch, Wm. W. *Employee aptitude survey: Technical Report.* Los Angeles: Psychological Services, Inc., 1963.

Stahl, O. G., & Staufenberger, R. A. (Eds.) *Police personnel administration,* Washington, D. C.: Police Foundation, 1974.

# APPENDIX A

# JOB PERFORMANCE DESCRIPTION BOOKLETS FOR PRECINCT PATROLMEN INVESTIGATIVE PERSONNEL SERGEANTS INTERMEDIATE COMMANDERS

JOB PERFORMANCE DESCRIPTION BOOKLET*

for

PRECINCT PATROLMEN

* Feel free to page through this booklet before the rating session begins. The directions are more complicated than the usual rating form -- however, once you get the idea, the ratings are easy to do and will go quickly.

PLEASE DO NOT START THE RATINGS UNTIL ALL DIRECTIONS HAVE BEEN GIVEN.

84

## JOB PERFORMANCE DESCRIPTION SCALES FOR PRECINCT PATROLMEN

Measuring job performance has always been an important issue for both supervisors and subordinates. Supervisors want to know how well a man is performing his job so that they can make decisions about him such as salary increases, promotions, etc. And all men, whether supervisors or subordinates, want to know how well they are doing their own job so that they can pay more attention to the things they do poorly and thereby improve.

However, measuring job performance is a difficult thing to do. Though we all speak English, we quite often have a different understanding of what some common words mean, such as good, poor, average, etc. Because we differ in what we mean by these words, we often cannot agree on whether a person's job performance is "good," "average," or "poor." Numbers don't help either, because people still have trouble agreeing on what they mean in terms of job performance. Another problem in measuring job performance is that most people find it hard to make up their minds about what the major characteristics are of what they want to measure.

Let's say you were going to paint a room brown for a friend. You show him the color of paint you have and he says that the color you have is "a lot less brown" than the color he wants. So you change the combination of colors and he says that the color you have is "not quite as brown" as the one he had in mind. You change the combination of colors that make up brown and he says, "That color is a little less brown than I wanted." It is clear that what your friend means by "a lot less brown," "not quite as brown," "a little too brown," etc. is not the same as what you mean by "a lot less brown," "not quite as brown," etc. This is an example of the problem people have with measuring things using words that are not closely related to the thing they are measuring, in this case the color brown.

If your friend had told you that the paint was too shiny, you would have known he was talking about something other than the color of the paint, but what if he took the shininess of the paint into account and didn't tell you? This is an example of having different ideas of what is involved in measuring brownness.

Well, you got smart and decided that you would make up a whole series of brown paint chips according to different combinations of colors and arrange them so that your friend could look at them and pick the color he expected his room to be painted in. Of course, not every possible shade of brown was included in the colors you showed him, but the examples you had mapped out formed a scale of brown.

We have tried to develop job performance rating scales for precinct patrolmen that consist of about the same thing. Each scale is a list of possible examples of behavior. The examples form an outline of what is involved in doing one aspect of the job of a patrolman. The examples also serve as guideposts or trail markers along the scale, marking off more and more effective performance.

We developed 11 different rating scales plus an overall performance scale for the job of precinct patrolman, based on information provided by patrolmen and sergeants in a major metropolitan city. The names of these scales are:

A. Crime Prevention
B. Using Force Appropriately
C. Traffic Maintenance and Control
D. Maintaining Public Safety and Giving First Aid
E. Investigating, Detecting, and Following up on Criminal Activity
F. Report Writing
G. Integrity and Professional Ethics
H. Dealing Constructively with the Public
I. Handling Domestic Disputes
J. Commitment, Dedication, Conscientiousness
K. Teamwork
L. Overall Job Performance

Instructions: In the booklet accompanying these directions you will find 12 job categories representing performance dimensions which are important for the job success of patrolmen. Each job category has immediately below it a definition of the category and seven or eight performance examples gathered from policemen in a number of cities. Each performance example "anchors" the portion of the scale opposite it. That is, on the job category, "Crime Prevention," the performance example which begins, "The officer went to every 24-hour gas station . . ." represents the kind of behavior or performance which should be rated at 8 or 9. Thus, the performance examples should help to define further the job category and should provide "benchmarks" for defining points on the nine-point scales.

In addition to the job categories booklet containing the 12 categories, their definitions, and the performance examples, we have provided you with a rating worksheet for each job category with the names of the patrolmen you will be rating on the left side of the worksheets. You will then use these worksheets to record the ratings you assign to your subordinates. When you are ready to begin the rating task, you should first read over the definition and the performance examples for Job Category A, "Crime Prevention." Get the content of this job category firmly in mind. Now consider the typical performance within this category of your first subordinate ratee. Compare his typical performance with the performance represented by the eight "benchmarks." Assign this man the rating which best typifies his level of performance in the Crime Prevention area compared with the example anchors. Record that rating opposite his name on

the worksheet entitled Crime Prevention. Then, go on to the next subordinate and rate him on the same job category. When you have finished rating all of your subordinates on Crime Prevention, turn to the next category (Job Category B, "Using Force Appropriately") and proceed the same way you did for the first category. Follow the same procedures for the other 11 categories as well.

Things to guard against: Several sources of error can contribute to inaccuracies in your ratings. Mentioning them briefly may help you to guard against making them. Here are suggestions for overcoming them:

1) Consider each performance category separately from all the rest. An almost universal error in ratings is called HALO ERROR. It occurs when the rater gives about the same ratings to a person on all aspects of performance. Usually this error occurs because a rater has not really taken enough time t get clearly in mind what each separate category of performance refers to. Remember we are asking you to describe or rate each of your subordinates on 12 categories of performance. As you consider each of the persons you are rating try to avoid getting into the habit of giving about the same rating to him on each job category. Consider each category separately from all others. Be sure to rate all your subordinates on each category before going on to the next category.

2) Consider each subordinate's performance over time and not on just one or two occasions. Another type of error occurs when a rater is influenced by just some single event or some recent occurrence. As you consider each subordinate's performance, think back over all the time you have known him and try to avoid being influenced by just one or two events. Base your ratings on all your observations of him and not just a few.

3) Avoid being misled by such things as appearance, education, family background, and other personal characteristics. Another common error in rating is called STEREOTYPE ERROR. It occurs when a rater allows himself to be influenced by other things than what the person has actually done on the job. In considering each subordinate's job performance, try to ignore everything else you may know about that person. Give your rating based strictly on what the individual has done on the job.

4) Avoid using your own definitions for the various job categories. A common reason for inaccurate ratings is because raters have different definitions of the job categories. Terms such as Crime Prevention, Maintaining Public Safety, and Dealing Constructively with the Public, etc. can have different meanings for different raters. This is why it is so very important for

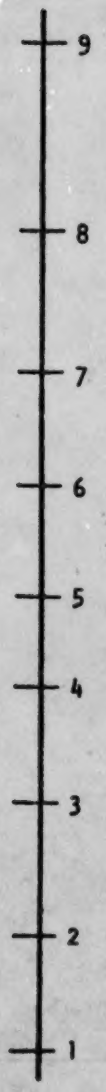
you to read the definitions and performance examples carefully for these job categories. Avoid any previous impressions of what these things have meant to you. Base your ratings on the definitions which are provided in the rating booklet you have received.

Remember, these ratings will be used solely for the purpose of validating the experimental battery of tests given earlier to those persons you are rating. Neither performance evaluations nor test scores will be shown to any member of the police department. Thus, you should feel free to be open and honest in completing these ratings. Thank you very much for your cooperation.

## CRIME PREVENTION

Job Category A

Knowledge of effective crime prevention, such as silent alarms, security fences, lighting and random patrol; educating citizens to aid in deterring criminal activity or in aiding apprehension of suspects; maintaining security in keeping relevant information from potential criminals; being aware of trends in criminal activity; keeping an eye on potential or known criminals in the area.



9 The officer went to every late night gas station in his area to alert the attendants about a group of hold-up men who had been hitting gas stations. He left a description of the men, a phone number to call and detailed instructions on what to do if the men were spotted. Because of his actions the hold-up men were apprehended.

8 When eight burglaries had occurred in a small area, the officer told a citizen that he would tell them how to help if they wished. The citizen organized a coffee party where the officer's tips on what to do led to the arrest of six young men.

7 An officer, after checking apartment house parking lots for car prowlers, would make a note of any apartment that didn't have good lighting and then tell the caretaker during the day.

6 The officer advised a bar owner who had been burglarized to wire a bell to the back door so a bartender who lived above the bar could tell when there was a break-in.

5 After there had been a rash of burglaries, the officer began spending more time patrolling the area they occurred in.

4 While on his night beat, an officer observed a business with one of the windows open. Finding no evidence of a break-in, he failed to report the open window to the owner the next day.

3 A burglar who was being transported to jail asked how the officer had known he had broken in. The officer then explained all about silent alarms--how they worked, how to spot them, etc.--educating him for his next job.

2 While on patrol an officer takes his coffee and lunch breaks at the same time and same place every night. He also patrols his beat in the same pattern every night.

1

## USING FORCE APPROPRIATELY

Job Category B

Keeping one's "cool" under pressure or personal abuse; being able to judge and to utilize the correct amount of force to resolve an incident promptly and effectively; avoiding acts that might be seen as brutal or sadistic; retaining composure when confronted with hostility and provocation.

9 A deskman calmly convinced a man who was pointing a rifle at him to hand it over rather than shooting the man when he had the chance.

8 In a fight with a traffic violator, the violator knocked one officer down, took his revolver, and shot six shots at the officer's partner, hitting him four times. The wounded officer pulled his revolver and drew a bead on the violator, who then threw the empty gun down and raised his hands. The wounded officer did not fire, but instead kept the violator covered until he was in custody.

7 An officer stopped a car for a traffic violation and the driver assaulted the officer with obscenities and verbal abuse. The officer wrote the tag and calmly explained why the man got the tag and how he could handle it, still amid a barrage of obscenities.

6 The officer grabbed the arm of a girl attacking her boyfriend with an ice pick, narrowly saving him. The officer was then assaulted by her, and had his shirt ripped by the ice pick before he struck her in the head with his gun to subdue her.

5 The officer waited for two young men who had been rowdy and noisy in a restaurant to come back to their car to pick them up. He took them to a dark area several blocks from their car, kicked them in the ass, and told them to walk back to their car. He also said that they should stay out of the area, because their kind weren't needed.

4 While taking a very hostile and belligerent man to jail, the officer purposely threw him against the wall.

3 The officer slapped a man who was pestering a bartender to sell him a drink after hours.

2 A man stopped after being chased at high speed. Even though the situation was in hand, an officer from a second squad which pulled up began beating the man.

1

90

## TRAFFIC MAINTENANCE AND CONTROL

**Job Category C**

Concern for and effective actions for maintaining motorized and pedestrian traffic safety; knowing traffic ordinances; exercising caution in apprehending speeders and other offenders; responding quickly to accidents and taking proper actions to preserve life and protect property; protecting the accident scene.

9 After an officer became aware that a dangerous intersection had no traffic control devices and that a high hedge was obstructing the view, he took it upon himself to contact the traffic engineers to have signs posted and the owner of the hedge to have it cut.

8 While tagging a driver for speeding in a school zone, the officer explained how unpredictable children were when playing and how much damage a car can do to a pedestrian.

7 An officer on routine patrol observed an emergency vehicle attempting to go through an intersection and immediately took measures to stop traffic and control the situation.

6 The officer tagged and towed a parked car which he found covered with snow, though it hadn't snowed for five days.

5 Investigating an accident, an officer used his squad car to block a street at the bottom of a hill. A car coming down the hill was unable to stop and hit the squad.

4 Observing a driver traveling at high speeds down a residential area late one night, an officer decided not to ticket the individual because the street was clear, but to warn him. [Even when the driver became impatient with being stopped, the officer gave only a warning.]

3 While on patrol, the squad car was almost involved in an accident with a car which turned right in front of it. The officer disregarded the partner's suggestion to give a tag and said, "No, I'll just chew them out."

2 While directing rush hour traffic from the middle of a very busy intersection, the officer began a needless conversation with a friend. He stopped directing traffic and stood with his friend in the middle of the street obstructing the flow and seriously endangering himself and his friend.

1

## MAINTAINING PUBLIC SAFETY AND GIVING FIRST AID

Job Category D

Concern for public safety; knowing and using the proper procedures for dealing with hazardous or emergency situations; evacuating and restricting activities in dangerous areas; giving quick and effective first aid when indicated.

9 At a propane gas tank leak, the officer requested cars to block specific intersections. He then shut down two nearby companies and began evacuating the area, all before receiving orders from his supervisor.

8 While watching a parade, an old man collapsed. An officer rushed up, pushed the crowd back, gave mouth-to-mouth resuscitation and saved the man's life.

7 Arriving at a house with two burning fire bombs on the front porch, the officer evacuated the house, contacted the fire department, and extinguished the flames with dirt.

6 Responding to a call about a burning car, an officer, noticing a fire near the gas tank, evacuated the area of bystanders and contacted the fire department.

5 In response to a suicide attempt where a girl had slashed her wrists, the officer administered proper first aid to stop the bleeding.

4 At a bomb threat to a business, the officer evacuated the building, but did not evacuate adjacent buildings.

3 An officer saw that the sidewalk next to a building that was being wrecked was not blocked off and that people might be hurt by debris, but he did nothing about it.

2 At an auto accident a victim complained that he was injured, though there were no signs of any injury. The officer told the person that he was faking and then refused to call an ambulance.

1

## INVESTIGATING, DETECTING, AND FOLLOWING UP ON CRIMINAL ACTIVITY

**Job Category E**

Being fully informed about all wanted felons; being alert to unusual circumstances or out-of-the-ordinary situations; protecting the crime scene to maintain the integrity of evidence; attentiveness to detail; effective questioning of witnesses; verifying both suspects' and witnesses' answers; following up on all clues or leads.

9 An officer was called to a domestic involving a man with a .38 caliber revolver and two companions trying to get an ADC check. Six hours later, when an armed robbery took place in another district by three men with a .38, he immediately provided detectives with names of suspects and a car description, leading to arrests and recovery of the loot.

8

7 After finding footprints leading up to a wall of a warehouse, but no prints leading away, an officer called for a dog and a key for the warehouse. A burglar who had scaled the wall and entered through a ventilator shaft was found inside.

6 At the scene of a man with a gun call, the officer found a gun which he handled carefully to preserve any fingerprints.

5 An officer requested a listing on a car parked in front of a house occupied by people who were suspected of associating with burglars. Since the plates listed to another car, it was towed away.

4 The officer was given a knife that appeared to have blood on it by a man who had found it laying in his yard. The officer put the knife in the glove compartment and forgot about it.

3 Early one morning an officer noticed a young man standing near a used car lot. When the officer asked him what he was doing, he said he was waiting for his dad to pick him up. The officer left and the next morning several car parts were reported

2 missing from the used car lot.

At the scene of a burglary where many TV sets were taken, the officer was told by a neighbor that he had observed a truck at the scene earlier in the evening. The officer failed to get

1 the neighbor's name and did not follow up the information.

## REPORT WRITING

Job Category F

Knowing and using the proper style or mode for reporting or communicating information; providing complete reports of one's actions; specifying all details which may aid in reconstructing the incident; using the correct grammar and language in reports.

9 The officer's report of a robbery of a person where a suspect was apprehended contained not only a standard account of the crime, but details of the weather and lighting conditions at the scene and a field sketch of the crime scene.

8 The officer's resume contained the names of all witnesses he questioned, their occupation, residence, phone number, and reason for being in the area.

7 An officer helped two other officers write a report of a felony arrest so that it contained all necessary information and was acceptable to the county attorney.

6 On a report form for a burglary of a dwelling, the officer filled in all of the spaces properly, but the body of the report was brief and he failed to explain some minor details.

5 An officer's offense report did not mention the type of residence burglarized.

4 An officer's report contained many incomplete sentences or fragments, such as "held suspect while partner opened case he was carrying."

3 An officer was called to a house burglary, investigated, and filed a report. The next day, he was requested by the Detective Division to return to the scene and redo his incomplete report.

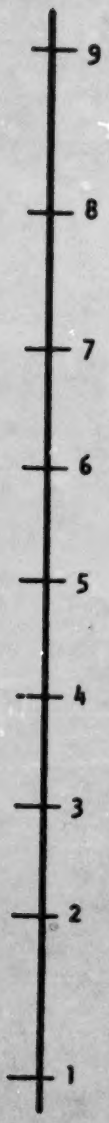
2 An officer wrote an incomplete and messy report of a serious crime. He was told to redo the report, but instead he simply filled in the missing information. (The report was again returned and the officer was told to retype it.)

1

## INTEGRITY AND PROFESSIONAL ETHICS

Job Category G

Avoiding opportunities to use one's badge, uniform, or authority for personal gain; refusing bribes, inventorying all evidence; presenting evidence accurately and completely; avoiding situations which might compromise one's honesty.



9 The officer gave a businessman he knew a ride home, because the man was drunk. The next day the officer received an envelope containing $200 from the businessman. The officer returned the money and explained that he took the man home because he was a friend and expected nothing for it.

8 A man offered to pay the officer if he wouldn't enforce prostitution laws so tightly in his area. The officer refused, sent a memo to the Morals Division and observed the man even closer in the future.

7 An officer who was having financial problems was offered a $100 bribe by a drunk driver, but he immediately refused the money and added attempted bribery to the charges.

6 The officer gave a ride to a man who had left his house as a result of a domestic. At his destination, the man offered the officer some money for his troubles, but the officer declined.

5 After arresting two men drinking illegally and confiscating their bottle, the officer refused the money offers of "winos" along the street for the liquor.

4 On the way home from work, an officer would stop at a drive-in that fed uniformed officers for nothing and pick up dinner for his family.

3 Two officers walked into a bar and one officer asked for a Christmas bottle for each. When his partner said, "Put mine back, I don't want it," this officer took both bottles.

2 Answering a call to a D.O.A., an officer told the bystanders in the apartment building to go back to their rooms, that he would handle everything. His partner asked why he was searching the apartment, and the officer replied, "You never know what you can find, especially money."

1

## DEALING CONSTRUCTIVELY WITH THE PUBLIC

Job Category H

Courtesy and understanding; helping citizens in matters that may not be strictly police business; maintaining and improving the police department's image in the eyes of the public; knowing about and using other agencies for referral of citizens who have special problems.

9 The officer made a service call to a destitute family. He called the proper authorities to obtain assistance for them and bought a tree and presents to make their Christmas happier.

8 Answering a call in which a blind man had been robbed of his grocery money, the officer went to a nearby church to collect food coupons and then took the blind man to the store and assisted him in buying groceries.

7 The officer had an elderly woman who had lost her house keys sit in the squad, out of the cold, while he gained entry. After she was inside the officer replaced the storm window he had removed.

6 Because the bartender admitted having served the man too much liquor, the officer didn't arrest the man who was slumped over the bar, but instead walked him home.

5 A girl's boyfriend was abusive toward an ambulance attendant, who spoke in a loud voice to the hysterical girl in an attempt to get through to her. An officer called the boyfriend aside and told him in no uncertain terms that he was wrong and to stop interfering.

4 A man flagged an officer down and asked if he could get a jump start since his car battery was dead. The officer said he wasn't allowed to and drove off.

3 The deskman was listening to a man's questions about a traffic accident when the phone rang. He just picked up the phone saying, "The forms are on the table," but he didn't answer the man's questions.

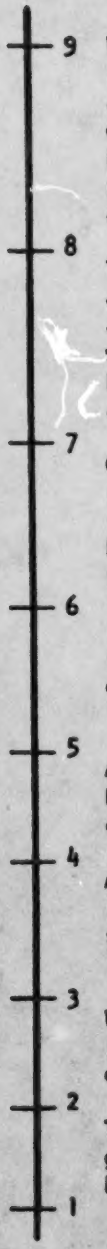
2 A depressed alcoholic committed suicide by jumping out of a hotel room after a minister left the room to get the man's bag. The investigating officer said, "When you left the room to go down to the car, did the guy tell you he'd meet you?"

1

## HANDLING DOMESTIC DISPUTES

**Job Category I**

Holding back and using restraint in working with arguments and fights between husbands and wives, boyfriends and girlfriends, or other domestic combatants; exercising caution; mediating between parties while maintaining impartiality; referring citizens to appropriate agencies for further help.



9 When the officer arrived at the scene of a domestic, he found that the husband had assaulted his wife but that she didn't want him arrested. She wanted to leave with her small children, so the officer helped the woman dress her children while he kept the husband in a separate room. While the officer drove them to her parents' home, he advised her of the various agencies that could assist her with her marital problems.

8 The husband and wife knew the officer by name and his first action at that domestic call was to get them in separate rooms, ask each what their problems were, and how he could help. The wife wanted the husband to go to bed or leave, since he had been drinking. The officer presented him with these alternatives and advised him to go to bed, which he did.

7 On the third call to a husband/wife domestic, the officer realized the wife was using the officer's presence to belittle the husband. The officer took her aside and told her to seek help for her problems, but that the officer would not allow her to misuse his authority.

6 In order to arrest a man without a fight, the officer at a domestic explained that by law he had to arrest the man, that he would call more officers if need be, and that the man might get hurt if he put up a fight.

5 At a domestic, the wife made a "citizen's arrest" of her husband. The son sided with his father, and the officer had a difficult struggle to subdue and jail him.

4 At a domestic, the officer advised the husband, who was drunk, to leave when his wife refused to sign a complaint. The domestic was settled, because the husband drove away, but he left under the influence of alcohol.

3 When the officer arrived at a domestic, the wife started to leave, but the officer called her back and the husband/wife domestic began again.

2 The officer took a gun away from a woman in a domestic, but gave it back to her before her husband had left, so that she had it reloaded as her husband was leaving.

1

## COMMITMENT, DEDICATION, AND CONSCIENTIOUSNESS

Job Category J

Exerting maximum effort at all times; responding to all calls; constantly updating and improving one's skills; being on time; readiness to provide police services at all times, both on and off duty; readiness to respond to needs at all times, both on and off duty; conducting oneself properly when off duty.

9 An off-duty police officer and his wife pulled into a gas station just after it had been held up. The officer told his wife to call the police, then gave chase on foot, apprehending one suspect.

8 An officer observed a person jump from a bridge into the Mississippi River. He jumped into the river and pulled the person to safety.

7 An officer recently assigned as deskman at the precinct received no instructions on what the job involved, so he read the job description and was able to handle all duties.

6 An officer occasionally assigned to a certain beat noticed juveniles hanging around a vacated building. The officer passed this information on to the men who were permanently assigned to the area.

5 An off-duty who was informed of a potentially dangerous situation; i.e., children digging into the side of a steep bank, failed to make note of it and did not remember to report it for several days.

4 The officer was in a cafe drinking coffee, even though he had told the dispatcher he was still at an accident.

3 After being informed at roll call every day two weeks prior to the date that officers were to change to winter uniforms, an officer came to work in his summer uniform.

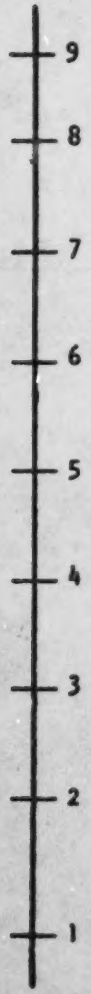
2 The officer was tipped off to a burglary, but got there too late because he took care of some personal business first.

1

## TEAMWORK

Job Category K

Having a good "feel" for what one's partner's actions are going to be without asking; backing him up and keeping his safety in mind at all times; keeping him informed; willingness to risk one's own safety to assure his protection; cooperating with other divisions or departments; assisting other law enforcement agencies such as the FBI or State's Attorney.



9 When the officer saw the criminal he and his partner had been tailing was about to shoot his partner, he yelled the criminal's name, which fouled his shot, saving the partner's life.

8

While two officers were closing in on a wanted criminal, the officer realized that his partner had not seen the gun the criminal had drawn. The officer yelled and alerted his

7 partner.

While searching some bars for a robbery suspect, one officer would stand in front of the bar while his partner would go

6 inside and look for the suspect.

When the officer received his days-off slip for the month, he called his partners and arranged the days off so that the days

5 off were acceptable to all.

At roll call an officer was given memos concerning his squad's work in a district. The officer put the information in his

4 pocket, failed to tell his partner what it concerned, and handled the situation by himself. Thus, his partner was unable to answer his supervisor's questions regarding the handling of the

3 instructions.

When asked to assist in arresting a drunk, the officer simply walked away, even though the drunk was being obviously

2 troublesome to his fellow officer.

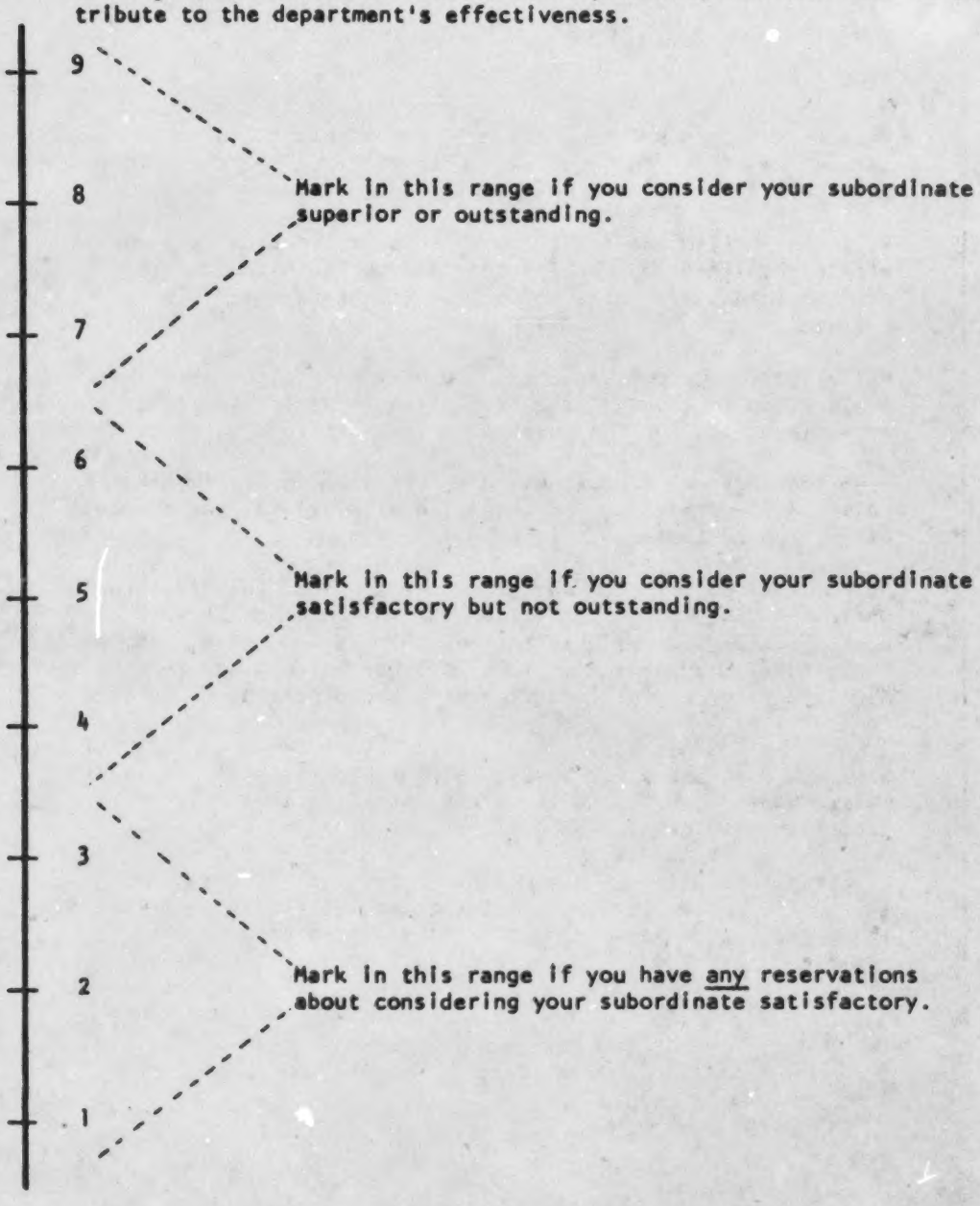
An officer remained in a squad car "apparently frozen" even though his partner got out of the squad, attempted to break up the fight, got involved in it, and was threatened by the

1 crowd.

## OVERALL JOB PERFORMANCE

Job Category L

Consider here the overall performance of a subordinate you are rating. The following should enter into your rating: how well does he do the total job; how valuable is he to the department; how much does he contribute to the department's effectiveness.


9
8
7
6
5
4
3
2
1
Mark in this range if you consider your subordinate superior or outstanding.
Mark in this range if you consider your subordinate satisfactory but not outstanding.
Mark in this range if you have any reservations about considering your subordinate satisfactory.

JOB PERFORMANCE DESCRIPTION BOOKLET*

f o r

INVESTIGATIVE PERSONNEL

* Feel free to page through this booklet before the rating session begins. The directions are more complicated than the usual rating form--however, once you get the idea, the ratings are easy to do and will go quickly.

PLEASE DO NOT START THE RATINGS UNTIL ALL DIRECTIONS HAVE BEEN GIVEN.

## JOB PERFORMANCE DESCRIPTION SCALES

## INVESTIGATIVE PERSONNEL

Measuring job performance has always been an important issue for both supervisors and subordinates. Supervisors want to know how well a man is performing his job so that they can make decisions about him such as salary increases, promotions, etc. And all men, whether supervisors or subordinates, want to know how well they are doing their own job so that they can pay more attention to the things they do poorly and thereby improve.

However, measuring job performance is a difficult thing to do. Though we all speak English, we quite often have a different understanding of what some common words mean, such as good, poor, average, etc. Because we differ in what we mean by these words, we often cannot agree on whether a person's job performance is "good," "average," or "poor." Numbers don't help either, because people still have trouble agreeing on what they mean in terms of job performance. Another problem in measuring job performance is that most people find it hard to make up their minds about what the major characteristics are of what they want to measure.

Let's say you were going to paint a room brown for a friend. You show him the shade of paint you have and he says that the shade you have is "a lot less brown" than the shade he wants. So you change the combination of colors and he says that the shade you have is "not quite as brown" as the one he had in mind. You change the combination of colors that make up brown and he says, "That shade is a little less brown than I wanted." It is clear that what your friend means by "a lot less brown," "not quite as brown," "a little too brown," etc. is not the same as what you mean by "a lot less brown," "not quite as brown," etc. This is an example of the problem people have with measuring things using words that are not closely related to the thing they are measuring, in this case the color brown.

If your friend had told you that the paint was too shiny, you would have known he was talking about something other than the shade of the paint, but what if he took the shininess of the paint into account and didn't tell you? This is an example of having different ideas of what is involved in measuring brownness.

Well, you got smart and decided that you would make up a whole series of brown paint chips according to different combinations of colors and arrange them so that your friend could look at them and pick the shade he expected his room to be painted in. Of course, not every possible shade of brown was included in the colors you showed him, but the examples you had mapped out formed a scale of brown.

We have tried to develop job performance rating scales for precinct patrolmen that consist of about the same thing. Each scale is a list of possible examples of behavior. The examples form an outline of what is involved in doing one aspect of the job of a patrolman. The examples also serve as guideposts or trail markers along the scale, marking off more and more effective performance.

We developed 12 different rating scales plus an overall performance scale for the job of investigator, based on information provided by investigators and detectives in a major metropolitan city. The names of these scales are:

A. Investigating the Scene of a Crime
B. Arrest, Search, and Seizure
C. Interrogating Suspects
D. Investigating a Case
E. Developing and Utilizing Informants
F. Report Writing and Paperwork
G. Appearing in Court
H. Public Relations
I. Dealing with Juveniles
J. Cooperating with Other Officers and Divisions
K. Conscientiousness and Dedication
L. Integrity and Honesty
M. Overall Job Performance

Instructions: In the booklet accompanying these directions you will find 13 job categories representing performance dimensions which are important for the job success of investigators. Each job category has immediately below it a definition of the category and seven or eight performance examples gathered from policemen in a number of cities. Each performance example "anchors" the portion of the scale opposite it. That is, on the job category, "Investigating the Scene of a Crime," the performance example which begins, "At the scene of a robbery where one suspect . . ." represents the kind of behavior or performance which should be rated at 8 or 9. Thus, the performance examples should help to define further the job category and should provide "benchmarks" for defining points on the nine-point scales.

In addition to the job categories booklet containing the 13 categories, their definitions, and the performance examples, we have provided you with a rating worksheet for each job category with the names of the investigators you will be rating on the left side of the worksheets. You will then use these worksheets to record the ratings you assign to your subordinates. When you are ready to begin the rating task, you should first read over the definition and the performance examples for Job Category A, "Investigating the Scene of a Crime." Get the content

of this job category firmly in mind. Now consider the typical performance within this category of your first subordinate ratee. Compare his typical performance with the performance represented by the eight "benchmarks." Assign this man the rating which best typifies his level of performance in the area of "Investigating the Scene of a Crime" compared with the example anchors. Record that rating opposite his name on the worksheet. Then, go on to the next subordinate and rate him on the same job category. When you have finished rating all of your subordinates on the first category, turn to the next category (Job Category B, "Arrest, Search, and Seizure") and proceed the same way you did for the first category. Follow the same procedures for the other 12 categories as well.

Things to guard against: Several sources of error can contribute to inaccuracies in your ratings. Mentioning them briefly may help you to guard against making them. Here are suggestions for overcoming them:

1) Consider each performance category separately from all the rest. An almost universal error in ratings is called HALO ERROR. It occurs when the rater gives about the same ratings to a person on all aspects of performance. Usually this error occurs because a rater has not really taken enough time to get clearly in mind what each separate category of performance refers to. Remember we are asking you to describe or rate each of your subordinates on 13 categories of performance. As you consider each of the persons you are rating try to avoid getting into the habit of giving about the same rating to him on each job category. Consider each category separately from all others. Be sure to rate all your subordinates on each category before going on to the next category.

2) Consider each subordinate's performance over time and not on just one or two occasions. Another type of error occurs when a rater is influenced by just some single event or some recent occurrence. As you consider each subordinate's performance, think back over all the time you have known him and try to avoid being influenced by just one or two events. Base your ratings on all your observations of him and not just a few.

3) Avoid being misled by such things as appearance, education, family background, and other personal characteristics. Another common error in rating is called STEREOTYPE ERROR. It occurs when a rater allows himself to be influenced by other things than what the person has actually done on the job. In considering each subordinate's job performance, try to ignore everything else you may know about that person. Give your rating based strictly on what the individual has done on the job.

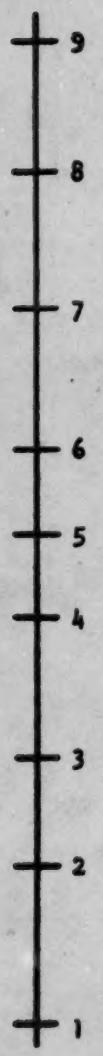
4) Avoid using your own definitions for the various job categories. A common reason for inaccurate ratings is because raters have different definitions of the job categories. Terms such as "Arrest, Search, and Seizure," "Interrogating Suspects," and "Public Relations," etc., can have different meanings for different raters. This is why it is so very important for you to read the definitions and performance examples carefully for these job categories. Avoid any previous impressions of what these things have meant to you. Base your ratings on the definitions which are provided in the rating booklet you have received.

Remember, these ratings will be used solely for the purpose of validating the experimental battery of tests given earlier to those persons you are rating. Neither performance evaluations nor test scores will be shown to any member of the police department. Thus, you should feel free to be open and honest in completing these ratings. Thank you very much for your cooperation.

## INVESTIGATING THE SCENE OF A CRIME

Job Category A

Coordinating the information search at the scene of a crime; supervising personnel at the scene; recognizing, collecting, preserving, and processing physical evidence; searching for and identifying witnesses; determining the nature of the crime and evaluating possible charges against suspects.



**9** At the scene of a robbery where one suspect was captured and the victim wounded, a homicide investigator delegated specific areas of responsibility to each of the patrolmen and other investigators at the scene.

**8** At the scene of a shooting, an investigator obtained information from the uniformed officers, had pictures taken of the scene and followed a trail of the victim's blood, taking samples along the way.

**7** While conducting a search in -20° weather for a homicide suspect, an investigator discovered a footprint in the snow, covered the print, and assisted the Bureau of Identification in taking photos of the print.

**6** At the scene of a shooting in which a rifle was used, the investigator searched the ground for spent casings.

**5** At a burglary scene, the investigator failed to have the photographer take a picture of the point of entry.

**4** Called to the scene of a dwelling burglary involving juvenile suspects, the investigator found a positively identifying fingerprint on one article, but did not examine other items handled by the suspects and did not take plaster casts of footprints.

**3** During the investigation of a shooting scene, the investigator failed to segregate the witnesses, prior to questioning and allowed people to enter the scene, thus contaminating evidence.

**2** The investigator called to the scene of a burglary of business did not make a written supplement to the original report, took no photographs, failed to inventory evidence, failed to search the immediate area of the crime, and did not note the business addresses of the witnesses.

**1**

## ARREST, SEARCH, AND SEIZURE

Job Category B

Making appropriate use of laws of arrest, search, and seizure; preparing search warrants; providing information for arrest warrants; using proper caution and force in arresting suspects and searching prisoners; advising suspects of their rights, conducting searches methodically; inventorying property.

9 When the investigator spotted a wanted felon on the street, he called for assistance, kept the suspect in sight until it arrived, and then told the uniform officers where to station themselves during the arrest.

8 Investigator found, seized and marked small pieces of wood found in the trunk of suspect's car. The wood chips were matched to drawers within a safe recovered at a later date.

7 While executing a search warrant for narcotics in a dwelling, the investigator recorded the serial numbers of various appliances, stereos, etc. he found in the dwelling. He later checked NCIC and the property room to see if the items were stolen.

6 The investigator obtained a search warrant prior to searching a suspect's car which was found near the scene of crime.

5 Investigator made a legal search with a warrant and found a gun believed to be stolen, but left same. After checking records, he ascertained the gun to be stolen, went back, and found that the gun had been removed.

4 While searching a house, the investigator did not account for the physical dimensions of the house and overlooked a hidden room concealed by a large cabinet.

3 Because the arresting investigator had handcuffed a check forger with his hands in front and failed to search the man, the forger was able to throw away stolen checks he was carrying.

2 An investigator picked up a murder suspect whose description had been broadcast, and took him to the scene for identification before searching him for weapons.

1

## INTERROGATING SUSPECTS

Job Category C

Planning and carrying out the interrogation of specific suspects; keeping an open mind during interrogation; recording confessions from suspects; confirming facts learned during interrogation.

9 An investigator was sent to another state to bring back a suspect wanted in Minneapolis. The suspect had been questioned by the FBI, local and state police to no avail. The investigator checked the man's background and found he had been a high school baseball star in Minneapolis. During the trip back, the investigator got the man to open up by using his old memories of baseball and Minneapolis as a bridge.

8 When the investigator found that the burglary suspect he was interviewing was antagonistic toward police, he asked the reasons why, and then adopted a sympathetic manner. This led to the confession and recovery of stolen property.

7 Two investigators tried to interview a suspect, but the suspect became belligerent and would not talk. One investigator left while the other continued to make small talk. Finally the investigator got the suspect to talk with him about the crime.

6 The investigator began his interrogation of a burglary suspect by saying, "Let's talk a little bit about your background so I can get to know you."

5 An investigator obtained a confession from a burglar suspect which closed the case, but he failed to find out what had been done with the stolen property.

4 The investigator began talking about snowmobiles (an interest he had in common with the suspect) in an interrogation of a robbery suspect. However, after an hour, the investigator and the suspect were still talking about snowmobiles and hadn't gotten around to the robbery.

3 After a long interrogation, a murder suspect dictated a statement. However, the investigator waited until the next day to have him sign the statement, at which time the suspect refused.

2 Because the evidence was insufficient for prosecution and an admission was needed, the officer told a burglary suspect that he was not advised of his constitutional rights because he had none.

1

## INVESTIGATING A CASE

Job Category D

Gathering and evaluating supporting evidence for prosecuting a case; following up clues; finding and interviewing witnesses; taking statements from victims and witnesses; identifying suspects by conducting line-ups, showing mug shots, and comparing fingerprints.

9 — The investigator located some stolen property, but only a fraction of the total loss from a big burglary. The investigator observed the suspected burglar's activities for over a week until he led him to the remainder of the stolen property.

8 — After an elderly woman who had been swindled out of $500 was unable to identify the suspects from mug shots, the investigator contacted cities in neighboring states and asked for photos and descriptions of confidence men. He then showed these new pictures to the woman.

7 — Checking on the license number of a car used in a burglary, the investigator found that the address and owner were non-existent. The investigator checked to see who the previous owner was and contacted him for information on the present owner.

6 — Even though the victim of a burglary had seen the suspects at a distance of 25 feet and was unsure of what they looked like the investigator showed the victim a lengthy series of mug shots.

5 — Because two hold-up suspects had worn ski masks during the crime, the investigator did not hold a line-up.

4 — An investigator charged two suspects arrested for stealing from parking meters with petty theft rather than felony theft of government funds.

3 — An investigator did not check the ownership of a car a robbery suspect was driving when arrested.

2 — Because the investigator did not like to go into the area a burglary had occurred in, he conducted his investigation over the telephone.

1

## DEVELOPING AND UTILIZING INFORMANTS

Job Category E

Locating and developing informants; controlling and protecting them; gathering and evaluating information from informants.

9 While testifying, the officer refused to identify an informant whose information led to the arrest of a gang of burglars because he had promised the informant that he would not be identified.

8 The investigator recovered stolen property for a shoddy character and used this to develop a friendship with the man which led to the man giving the investigator information about criminals he knew.

7 An investigator worked at developing informants on a 24-hour basis by encouraging criminal types he met to phone his home or office whenever they needed assistance.

6 When a morals squad investigator broke up a "crap" game, he recognized one of the participants as an ex-con on parole. The investigator told the ex-con that if he provided the robbery division with information on a series of recent armed robberies, the investigator would not contact the man's parole officer.

5 When interviewing an informant about a prospective burglary, the investigator did not ask the informant how he happened to know about the planned burglary or his reasons for cooperating with the police.

4 An investigator in Morals arrested a prostitute and refused to trade her release for information on "hard" narcotics, even though he knew the information was probably good.

3 A morals investigator, while attempting to develop a woman as an informant, became intimately involved with her.

2 While interviewing a suspect arrested based on information supplied by an informant, the investigator revealed to the suspect the identity of the informant.

1

## REPORT WRITING AND PAPERWORK

**Job Category F**

Reporting all necessary information concisely; presenting information in chronological sequence; using correct terminology in reports; writing memos and instructions for other investigators; keeping proper records and files.

9 — In describing a DOA, the investigator wrote so complete a report that when a question arose as to what kind of watch the suspect was wearing, the information was found in the investigator's report.

8 — On a lengthy case involving many investigators, this investigator summarized the individual reports, attached his summary to them and labeled it as a synopsis.

7 — When the investigator filled out supplementary reports, he made a point of filling in all of the blanks and then double checking all reports on the case to make sure there were no discrepancies.

6 — The investigator indicated on his reports the times witnesses could be reached for follow-up purposes.

5 — The investigator took complete notes at the scene. However, they included notes on matters which had no bearing on the crime.

4 — The investigator takes notes on a case on small slips of paper which he keeps in no apparent order in a box on his desk.

3 — An officer who made out an initial report in an assault case involving juveniles failed to get the ages, telephone numbers, schools attended, etc. of the witnesses and suspects.

2 — An investigator wrote a report on a homicide and in the report specified information as fact when in reality the facts were only assumptions.

1

## APPEARING IN COURT

Job Category G

Preparing information for testimony in court and giving testimony; reviewing reports; presenting information concisely and accurately when testifying; using proper language; working cooperatively with prosecuting attorney.

9 — In the trial of a child murderer, the officer on the witness stand did not exhibit his personal feelings against the defendant.

8 — While testifying in court, the investigator did not volunteer information and tried to answer questions concisely.

7 — Prior to testifying in court on a case which involved a great deal of physical evidence and a lengthy interrogation of the suspect, the investigator pulled the case from the file and reviewed it.

6 — An investigator brought his own report of a robbery and the witness reports with him when he went to testify in court.

5 — The officer makes a late appearance at a trial, thereby inconveniencing others involved.

4 — When the investigator was asked by the defense attorney how a murder suspect appeared when questioned, he said, "confused", when he actually meant to say "evasive and devious".

3 — During the trial of a burglary suspect, the investigator on the case gave conflicting and inadequate testimony, because he had not reviewed his reports.

2 — An officer lost his temper, swore at the judge, and stomped out of the court room after hearing a judge's decision.

1

## PUBLIC RELATIONS

Job Category H

Dealing constructively with the public; being courteous, understanding, and professional in handling suspects, witnesses, victims, and the public in general; maintaining and improving the department's image in the eyes of the public.

9 The newspaper printed a story about an investigator who had arranged with the Humane Society for a new dog for some children whose dog had been clubbed to death by a burglar.

8 A burglary investigator alerted all of the victims of cases he was working on to the status of his investigation and the results of any prosecution of suspects.

7 The investigator, a regular patron of a small grocery store, was informed by the owner that several small checks written by the same person had "bounced". The investigator contacted the man and convinced him to make his checks good.

6 An investigator told a woman who reported that a gang of criminals were monitoring her by radio waves that he had arrested them on orders from the FCC and that the criminals would be in jail for 20 years.

5 When a burglary investigator found out that an attorney who wanted his clients released to his custody for the holidays had also told his clients to say nothing, he refused to release them.

4 An investigator showed up exactly at 6:00 p.m., even though he had been told by the witness of a burglary that he didn't want to be interviewed until later in the evening.

3 The investigator told the mother of a child which had been molested, "As long as you live in this neighborhood, you'll have to expect such things."

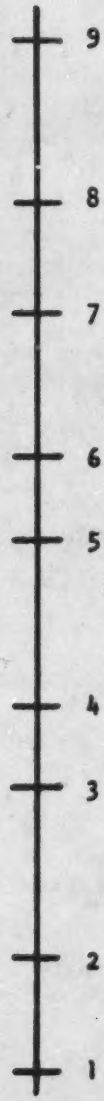
2 An investigator lost his cool and physically ejected a mentally retarded man who had come in to make a complaint.

1

## DEALING WITH JUVENILES

Job Category I

Contacting parents; talking and lecturing to juveniles in schools; deciding whether or not juvenile offenders should be charged; setting up conferences with juveniles' parents and victims; referring juveniles to proper agencies; following proper procedures when interviewing and interrogating juveniles.



9 When a youth gets into trouble, the juvenile officer tries to work out a solution by first talking with the parents and then setting up family counseling or getting help from Big Brothers, Big Sisters, Boys Club, etc. Only after everything else fails does he refer the youth to juvenile court.

8 The juvenile officer got the parents of some juveniles to consent to have their boys clean a garage they had written obscene words on.

7 A juvenile officer stationed in a school made up a full wall display in his office of narcotics posters, books, pamphlets, accessories, articles, etc. Students would come to view the display and he would talk with them about drugs.

6 A juvenile officer calls former delinquents he has worked with to see how they are doing.

5 When the investigator could not get a youth to admit to a series of garage burglaries in which the youth was the prime suspect, the officer called the youth's parents, told them his suspicions, and warned them to keep better tabs on the boy, or he was going to get caught.

4 A juvenile officer told a group of youths that anyone who took drugs was a criminal and should be treated like one.

3 An investigator assigned a case of minor property damage due to juveniles contacted the suspects and advised them to stop. However, he did not advise them to pay the victim for the damage, he did not contact the parents, and he did not contact the victim to advise him of the results of the investigation.

2 While standing in the playground after school, a fifth grader taunted and swore at a juvenile officer. The officer then grabbed the youth and slapped him.

1

## COOPERATING WITH OTHER OFFICERS AND DIVISIONS

Job Category J

Working cooperatively with other investigative divisions; keeping others informed of matters concerning them; assisting other departments and agencies such as sheriff or FBI.

9 When the investigator learned that some men were planning a burglary in an adjoining state, he put the men under surveillance and sent all of the information he had to the concerned police agencies in the adjoining state.

8 Having learned from an informant that a juvenile was implicated with an adult in four burglaries, the juvenile investigator solicited the aid of, and cooperated with, a burglary investigator in obtaining a search warrant and investigating the case.

7 After uniformed officers made an arrest of a felony suspect, the investigator continued to give the officers complete details on the progress of the investigation.

6 When a juvenile officer determined a youth had burglarized a home and stolen a pistol along with other property, he sent a description of the youth to the uniformed officers in the precinct with a request for a "pick-up".

5 A narcotics investigator obtained information on a possible burglary suspect. He checked the info out and solved the case without contacting the Burglary Division.

4 The investigator ordered two uniformed officers to keep a certain house under surveillance and to report any activity. However, he neglected to explain the general purpose of the investigation.

3 An investigator who was assigned a case involving the theft of Federal funds, refused to share information with the FBI, because he was afraid they would solve the case first.

2 An investigator did not share information concerning a possible robbery of a bar, even though he could have contacted the Robbery Division. He planted himself in the bar hoping he would apprehend the suspects, because he wanted the glory of capturing them by himself.

1

## CONSCIENTIOUSNESS AND DEDICATION

Job Category K

Exerting maximum effort at all times; attempting to do one's best on all cases; constantly updating and improving one's skills; being interested in all aspects of police work; working over-time willingly; providing police service both on and off duty.

The investigator reviewed a year old unsolved homicide case which had not been assigned to him. He re-examined all the information and developed new leads which resulted in solving the case.

An investigator stayed many hours after quitting time to complete his preliminary investigation of a robbery.

When an investigator was transferred to a new division, he checked all the statutes applicable to his new job, checked books out of the library on the new job, and reviewed completed cases from the files to familiarize himself with the new job.

An investigator was transferred from Robbery to Theft Division, but often receives information on robberies which he forwards to the Robbery Division.

The investigator attended college part-time and would sometimes read his text books at work when the load was light.

Because of his physical condition, the overweight investigator did not chase a suspect seen running from the scene of a robbery.

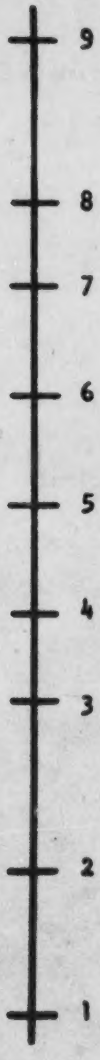
Because a case required a lot of leg work, the investigator stalled working on it until the shift changed and it was re-assigned.

An investigator called in sick on a Monday even though he wasn't, to try to avoid getting stuck with a heavy case load.

## INTEGRITY AND HONESTY

Job Category L

Treating all cases equally, not using the badge for personal gain; not letting personal prejudices or biases influence one's professional opinion; not accepting bribes or gratuities.




9 — An investigator was the first man on the scene of an attempted safe job. The burglar had left in such a hurry that $1500 was still in the safe. The investigator did not touch the money, even though he could easily have taken it and blamed the burglars.

8 — Even though an investigator had the opportunity to shade the truth to "get" a suspect whom he hated, he would not do it.

7 — A morals investigator was offered a $50 payoff each week if he would ignore a prostitute. The investigator took the bribe, and turned it in with a complete report to his supervisor.

6 — The investigator told a man who wanted to reward him for solving a robbery of his store to donate the money to the Policemen's Benevolent Fund if he wanted to reward him.

5 — An investigator spent more time and energy investigating a case in which his friend was the victim than he would have usually spent.

4 — An investigator accepted a good deal on a car as a reward for solving a theft case involving a used car lot.

3 — An investigator who frequently checked a bar where prostitutes worked, mentioned to the bar owner the trouble he was having with his car. The place the owner sent him to to get his car repaired never sent the investigator a bill and he never checked back to get one.

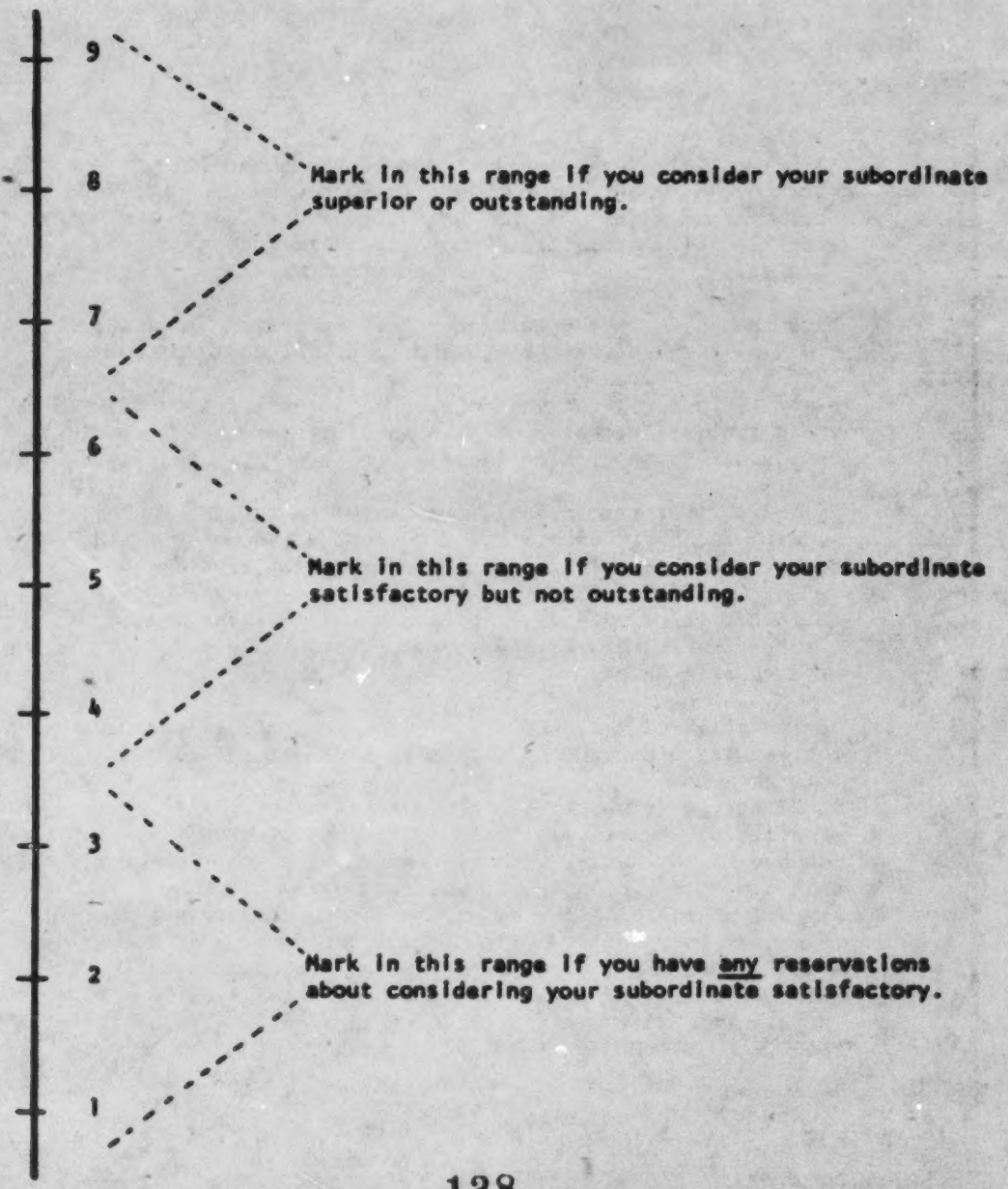
2 — After conducting a burglary investigation, the investigator (who was still on duty) offered to check the security of the victim's store; e.g., locks, lights, alarms, etc. for $150.

1

## OVERALL JOB PERFORMANCE

Job Category M

Consider here the overall performance of a subordinate you are rating. The following should enter into your rating: how well does he do the total job; how valuable is he to the department; how much does he contribute to the department's effectiveness.


9
8
7
6
5
4
3
2
1
Mark in this range if you consider your subordinate superior or outstanding.
Mark in this range if you consider your subordinate satisfactory but not outstanding.
Mark in this range if you have any reservations about considering your subordinate satisfactory.

JOB PERFORMANCE DESCRIPTION BOOKLET*

f o r

SERGEANTS

* Feel free to page through this booklet before the rating session begins. The directions are more complicated than the usual rating form--however, once you get the idea, the ratings are easy to do and will go quickly.

PLEASE DO NOT START THE RATINGS UNTIL ALL DIRECTIONS HAVE BEEN GIVEN.

## JOB PERFORMANCE DESCRIPTION SCALES

### SERGEANTS

Measuring job performance has always been an important issue for both supervisors and subordinates. Supervisors want to know how well a man is performing his job so that they can make decisions about him such as salary increases, promotions, etc. And all men, whether supervisors or subordinates, want to know how well they are doing their own job so that they can pay more attention to the things they do poorly and thereby improve.

However, measuring job performance is a difficult thing to do. Though we all speak English, we quite often have a different understanding of what some common words mean, such as good, poor, average, etc. Because we differ in what we mean by these words, we often cannot agree on whether a person's job performance is "good," "average," or "poor." Numbers don't help either, because people still have trouble agreeing on what they mean in terms of job performance. Another problem in measuring job performance is that most people find it hard to make up their minds about what the major characteristics are of what they want to measure.

Let's say you were going to paint a room brown for a friend. You show him the shade of paint you have and he says that the shade you have is "a lot less brown" than the shade he wants. So you change the combination of colors and he says that the shade you have is "not quite as brown" as the one he had in mind. You change the combination of colors that make up brown and he says, "That shade is a little less brown than I wanted." It is clear that what your friend means by "a lot less brown," "not quite as brown," "a little too brown," etc. is not the same as what you mean by "a lot less brown," "not quite as brown," etc. This is an example of the problem people have with measuring things using words that are not closely related to the thing they are measuring, in this case the color brown.

If your friend had told you that the paint was too shiny, you would have known he was talking about something other than the shade of the paint, but what if he took the shininess of the paint into account and didn't tell you? This is an example of having different ideas of what is involved in measuring brownness.

Well, you got smart and decided that you would make up a whole series of brown paint chips according to different combinations of colors and arrange them so that your friend could look at them and pick the shade he expected his room to be painted in. Of course, not every possible shade of brown was included in the colors you showed him, but the examples you had mapped out formed a scale of brown.

We have tried to develop job performance rating scales for sergeants that consist of about the same thing. Each scale is a list of possible examples of behavior. The examples form an outline of what is involved in doing one aspect of the job of a sergeant. The examples also serve as guideposts or trail markers along the scale, marking off more and more effective performance.

We developed eight different rating scales plus an overall performance scale for the job of sergeant, based on information provided by sergeants and their supervisors in a major metropolitan city. The names of these scales are:

A. Concern for Subordinates
B. Scheduling, Coordination, Deployment, and Manpower Allocation
C. Supervision
D. Performing Administrative and Inspection Functions
E. Decision Making and Initiative Where No Firm Guidelines Exist
F. Training and Planning
G. Integrity, Dedication, and Conscientiousness
H. Dealing Effectively with the Public and Superiors
I. Overall Job Performance

Instructions: In the booklet accompanying these directions you will find nine job categories representing performance dimensions which are important for the job success of investigators. Each job category has immediately below it a definition of the category and eight performance examples gathered from policemen in a number of cities. Each performance example "anchors" the portion of the scale opposite it. That is, on the job category, "Concern for Subordinates" the performance example which begins, "A patrolman was scheduled for . . ." represents the kind of behavior or performance which should be rated at eight or nine. Thus, the performance examples should help to define further the job category and should provide "benchmarks" for defining points on the nine-point scales.

In addition to the job categories booklet containing the nine categories, their definitions, and the performance examples, we have provided you with a rating worksheet for each job category with the names of the sergeants you will be rating on the left side of the worksheets. You will then use these worksheets to record the ratings you assign to your subordinates. When you are ready to begin the rating task, you should first read over the definition and the performance examples for Job Category A, "Concern for Subordinates." Get the content of this job category firmly in mind. Now consider the typical performance within this category of your first subordinate ratee. Compare his typical performance with the performance represented by the eight "benchmarks." Assign this man the rating which best typifies his level of performance in the area of "Concern for

Subordinates" compared with the example anchors. Record that rating opposite his name on the worksheet. Then, go on to the next subordinate and rate him on the same job category. When you have finished rating all of your subordinates on the first category, turn to the next category (Job Category B, "Scheduling, Coordination, Deployment, and Manpower Allocation") and proceed the same way you did for the first category. Follow the same procedures for the other categories as well.

Things to guard against: Several sources of error can contribute to inaccuracies in your ratings. Mentioning them briefly may help you to guard against making them. Here are suggestions for overcoming them:

1) Consider each performance category separately from all the rest. An almost universal error in ratings is called HALO ERROR. It occurs when the rater gives about the same ratings to a person on all aspects of performance. Usually this error occurs because a rater has not really taken enough time to get clearly in mind what each separate category of performance refers to. Remember we are asking you to describe or rate each of your subordinates on nine categories of performance. As you consider each of the persons you are rating try to avoid getting into the habit of giving about the same rating to him on each job category. Consider each category separately from all others. Be sure to rate all your subordinates on each category before going on to the next category.

2) Consider each subordinate's performance over time and not on just one or two occasions. Another type of error occurs when a rater is influenced by just some single event or some recent occurrence. As you consider each subordinate's performance, think back over all the time you have known him and try to avoid being influenced by just one or two events. Base your ratings on all your observations of him and not just a few.

3) Avoid being misled by such things as appearance, education, family background, and other personal characteristics. Another common error in rating is called STEREOTYPE ERROR. It occurs when a rater allows himself to be influenced by other things than what the person has actually done on the job. In considering each subordinate's job performance, try to ignore everything else you may know about that person. Give your rating based strictly on what the individual has done on the job.

4) Avoid using your own definitions for the various job categories. A common reason for inaccurate ratings is because raters have different definitions of the job categories. Terms such as

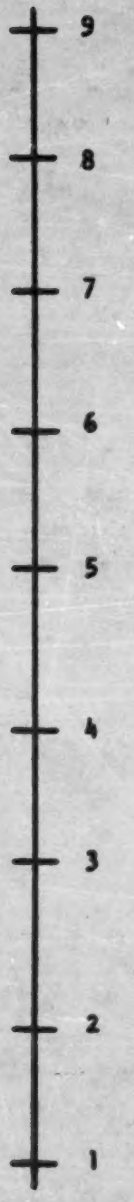
"Supervision," "Training and Planning," and "Concern for Subordinates," etc., can have different meanings for different raters. This is why it is so very important for you to read the definitions and performance examples carefully for these job categories. Avoid any previous impressions of what these things have meant to you. Base your ratings on the definitions which are provided in the rating booklet you have received

Remember, these ratings will be used solely for the purpose of validating the experimental battery of tests given earlier to those persons you are rating. Neither performance evaluations nor test scores will be shown to any member of the police department. Thus, you should feel free to be open and honest in completing these ratings. Thank you very much for your cooperation.

## CONCERN FOR SUBORDINATES

**Job Category A**

Showing concern for subordinates; being considerate; taking personal interest in their problems; giving recognition when deserved; using human relations principles; filling in for subordinates who have special problems; being aware of subordinates' needs and feelings.



**9** — A patrolman was scheduled for a vacation that would coincide with his wife's operation. When the operation was changed to a later date, the sergeant checked the work load and rescheduled his vacation to meet the change.

**8** — The sergeant wrote a letter of commendation to an officer who had done an outstanding job of clearing a case. He also sent a copy to the superintendent and posted a copy on the squad room bulletin board.

**7** — The sergeant told a man who showed up for roll call with a bad cough, sore throat, and runny nose to go home, because as he said, "I would rather you stayed home for a couple of days with a cold than work and get even sicker."

**6** — The sergeant told an officer who had come to him for some advice about personal problems that he (the sergeant) couldn't give him much help and maybe the officer should see a professional counselor.

**5** — The sergeant talked to an officer who had been having clashes with the other men on the shift and learned that he was frustrated because he hadn't made promotion. The sergeant suggested that the officer be a little less aggressive in the future.

**4** — When two of the three men on a squad complained of a "personality conflict" with their new partner, the sergeant, without further inquiry, placed the new man on a beat. The sergeant also did not discuss the reason for the change with the new man.

**3** — The sergeant told a patrolman who wanted to be excused two hours early from his tour to begin a long vacation trip to check back later. When the officer checked back, the sergeant told him to check back again. The tour ended with the officer never having received a yes or no answer.

**2** — When a patrolman had a personal problem and wanted to go home early to clear up the situation, the sergeant yelled and screamed and shouted obscenities at the patrolman in the presence of several other officers.

**1**

## SCHEDULING, COORDINATION, DEPLOYMENT, AND MANPOWER ALLOCATION

**Job Category B**

Matching manpower assignments with crime trends; assigning squads or unmarked cars to patrol certain areas; scheduling to take sick leave, vacation requests, and shift changes into account; deploying men using superiors' guidelines; scheduling without playing favorites.

**9** When uniformed patrols were unable to do much about an increase in burglaries in a residential area, the sergeant rescheduled his men and had several men come to work in plainclothes and patrol in unmarked cars until they were able to apprehend the burglars.

**7** The sergeant scheduled his men so that a maximum number were working on heavy "call load" days (i.e., Fridays and Saturdays) and a fewer number on lighter days (i.e., Sundays and Mondays).

**6** Because of complaints that officers were not being scheduled to work when they were really needed, the sergeant reviewed and posted the station policy on manpower needs and took over the job of reviewing days off requests and tried to work them in with station needs.

**5** Two men requested the same day off. The sergeant explained to them that the schedule would allow only one man to be off. After listening to their reasons for wanting the day off, the sergeant gave the day off to the officer who he felt had the greatest need.

**4** When unforeseen circumstances caused a shortage of regular shift personnel, the sergeant pulled men from the power shift to fill the vacancies.

**3** When the shift in the phone complaint room was short one man, the sergeant requested a replacement from another division rather than take the shift's relief man from the dispatcher section.

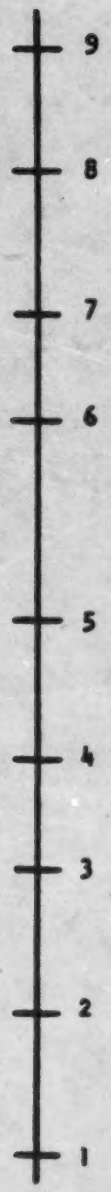
**2** When two men put in for the same day off, the sergeant disapproved both applications so he wouldn't have to decide which man had the better reasons for wanting the time off.

**1** Because the sergeant was a "good guy", he approved all vacation and days off requests without worrying about finding replacements.

## SUPERVISION

Job Category C

Supervising subordinates where guidelines or standard operating procedures exist; observing officers in the field to make sure they use proper procedures; warning and disciplining subordinates who are "out of line"; routine supervision at crime scene; evaluating subordinates.




Arriving at the scene of a burglary involving considerable loss, the sergeant assigned officers to protect the area from becoming contaminated, called for the investigative unit, and executed a preliminary search for evidence and witnesses.

When days off were cancelled and one man said he was going to be "sick" so he could take care of his part-time job, the sergeant warned the man that if this was the case, he had better be at home because he would be checked on.

When the sergeant heard a squad being sent to a call, he immediately drove to the scene and observed the squad handle the call.

After seeing an empty squad in front of a theater almost every night, the sergeant finally went in and found two patrolmen talking with the manager. When the officers told the sergeant they were just checking the theater, he said that was not their job and told them to stay away unless they received a call to go there.

When the sergeant arrived at the scene of an accident, he found that the patrolman was confused because it was his first accident investigation. The sergeant assigned him to direct traffic and took over the investigation himself.

The sergeant told some of his men that he had a list of the men on his shift who were "goof-offs", "slackers", and "bums", and that he was paying special attention to them.

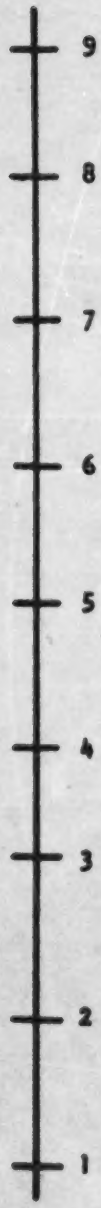
A desk sergeant received a call from a patrolman asking for advice on a burglary with a large loss. The sergeant listened and told the patrolman, "You're on the scene, you decide what to do."

The sergeant told his men on two different occasions to pick up their mail, but they did not respond. The sergeant then shrugged his shoulders and said he would do it.

## PERFORMING ADMINISTRATIVE AND INSPECTION FUNCTIONS

Job Category D

Keeping records and processing paper work; inspecting men and vehicles; critiquing and writing reports.




The sergeant went over a complaint report submitted by a patrolman, made the necessary changes, explained to the patrolman why the changes had to be made, and asked him to redo the report.

When the sergeant noticed that some officers were meeting departmental standards for appearance and equipment maintenance only on inspection days, he began holding impromptu inspections.

The sergeant made a daily inspection of all patrol cars, checking every area of the cars--the trunk, under the front and rear seats, etc.--and explained to the men the importance of proper inspection.

When the sergeant was assigned to find missing radio pack sets, he posted a sign over the pack sets, stating that each would be signed out through the desk officer, who was to report any unaccounted sets to the sergeant.

The sergeant warned men who came to work in uniforms that were not up to standards that the situation should be corrected before the next inspection. He repeated this warning at each inspection.

Because the sergeant disliked going over officers' reports for correct classifications, content, etc., he turned this responsibility over to a patrolman.

The sergeant sent a memo to the inspector, but the inspector never received it. Because he had made no copies and kept no notes, he could not remember all the details. He then wrote another memo which was considerably less detailed than the first.

When the sergeant was assigned to carry out a yearly inspection of the department, he stated, "Why? It won't do any good anyway," copied the last year's report, and forgot about the matter.

## DECISION MAKING AND INITIATIVE WHERE NO FIRM GUIDELINES EXIST

Job Category E

Making decisions where no guidelines or standard operating procedures exist; showing initiative; knowing when to get personally involved and when to delegate; reacting to unique or unusual emergency situations; showing leadership.

**9** A sergeant was at the scene of a dance when the crowd became unruly and began destroying property. He called in available squads, formed them up in a nearby parking lot, got the inspector's permission to lead the men in a sweep of the sidewalk, which dispersed the crowd, and then had squads patrol the immediate area.

**8** When the sergeant arrived at the scene of a bar fight, he found that two women who were being restrained by officers were creating a loud disturbance which was drawing a crowd. The sergeant immediately had the women removed from the area and then waited for more squads to help remove six men who were also arrested.

**7** When officers attempted to arrest a juvenile for possible rape charges, his mother physically sat on him and refused to move, claiming discrimination. The sergeant, noting the crowd of friends accumulating, called for more squads, and after attempting to persuade the mother to move, the suspect was forcibly removed while other officers formed a physical barrier.

**6** When the sergeant arrived at the scene of an auto accident and saw a car had gone through a guard rail and plunged 30 feet to the ground, he called the ambulance, the rescue squad, and a tow truck.

**5** The sergeant at the scene of a residential burglary in progress call found all the doors locked and no apparent sign of entry, though a neighbor said there was a light on that shouldn't have been on. He removed a screen, climbed into the house, and unlocked the doors to allow officers to search the house.

**4** While conducting a search for a wanted felon, a sergeant observed a group of militants forming. He assembled his men into a riot formation and immediately began dispersing the crowd.

**3** A sergeant frequently made a special effort to drive by a local hangout to tag the cars illegally parked in front.

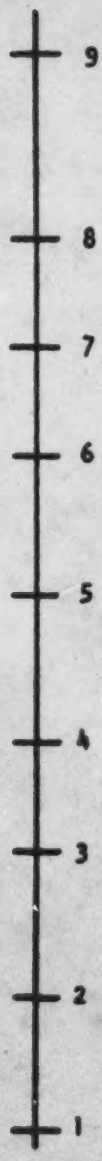
**2** Before he arrived at the scene of a man with a gun call (without receiving further information), the sergeant had the dispatcher send the special squad, gas, ambulance, dog squad and the inspector to the scene.

**1**

## TRAINING AND PLANNING

Job Category F

Orienting and training new recruits; thinking ahead so problems can be avoided; participating effectively in staff planning meetings; planning for training; anticipating problems; analyzing officers' weaknesses and suggesting training.




A patrolman was unsure he handled the arrest of a traffic violator correctly, so he asked the sergeant for advise. The sergeant explained the alternative ways the officer could have handled the situation and then discussed whether or not the officer had handled it correctly and how he would handle it if faced with the same situation again.

A sergeant was reinstructing his men in the proper handling of a shotgun. He had each man unload and load, and then set up a shotgun just as it is to be found in the squad boot.

Whenever the sergeant came across an article in a police magazine that he felt would benefit his patrolmen, he read it at roll call and commented on its interesting points.

When the sergeant saw that the back-up squads were not responding quickly and that this was a station-wide problem, he requested, at a staff meeting, that a station policy be set up.

When the sergeant was informed by the lieutenant that his men were consistently leaving important details out of their reports, the sergeant mentioned the matter in passing at roll call, but didn't follow through.

When a new patrolman was assigned to the precinct, the sergeant introduced him to the other officers and then assigned him to a squad with a poor performer.

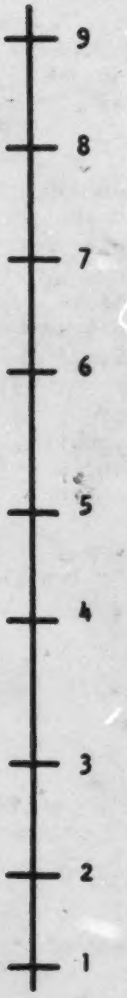
A sergeant was to present a roll call training session based on a slide projector presentation, but he knew little of its content because he had not bothered to check it out completely prior to roll call.

A sergeant never made any positive inputs during the monthly staff meetings and it was obvious that he didn't want to get involved in discussing problems and in establishing goals and priorities.

## INTEGRITY, DEDICATION, AND CONSCIENTIOUSNESS

Job Category G

Refusing bribes and special favors; setting an example for patrolmen; self-development; working hard; showing a good attitude; behaving properly on and off duty.




9 — The sergeant came to work 10 to 20 minutes early each day to exchange pertinent information with the previous shift sergeant and clear up any work that was left at shift's end.

8 — A sergeant asked to be transferred to a difficult assignment to learn what he could in this position, and he did a good job on the new assignment.

7 — A sergeant on patrol responded as quickly as possible to a robbery of business call and tried to apprehend the robbers in the area near where the incident had occurred.

6 — While patroling, a sergeant observed a car with out-of-state license plates parked in an alley behind the house of people who had police records. When he checked on the car and found that it was stolen, he apprehended the driver.

5 — A sergeant accepted some freshly caught fish from an officer. The sergeant gave the man the weekend off to go fishing as a favor, even though the shift was short.

4 — A sergeant had instructed his men to enforce one hour parking in a business district, but later told his men to "lay off" one block in the area, because a firm which gave police officers discounts had complained.

3 — Even though the sergeant told his men to use caution and good sense when driving, he was often seen driving very carelessly while on patrol.

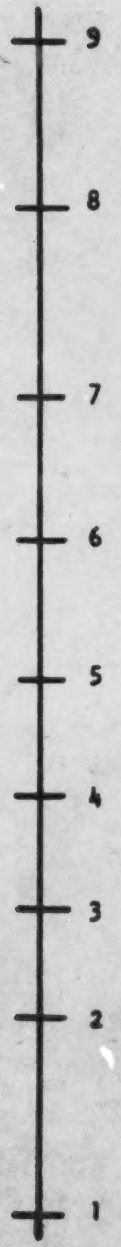
2 — While investigating the scene of a burglary, a sergeant stole three cartons of cigarettes.

1

## DEALING EFFECTIVELY WITH THE PUBLIC

Job Category H

Mediating between patrolmen and the public; dealing courteously and constructively with citizen and community problems; dealing effectively with superiors; awareness of department politics.



9

The sergeant kept checking the possible identity of a boy who had been run over and killed by a bus until he found the boy's mother. He stayed with the mother until the father came home. He then had a neighbor stay with the mother while he took the father to identify the boy.

8

A sergeant received a memo from an alderman about numerous accidents at a particular intersection. By checking through the Traffic Engineer's Office, he found that there had been only four accidents in a three-year period at the intersection. The sergeant wrote the alderman a letter explaining why added police enforcement was denied.

7

When the sergeant was encouraged by businessmen to assign two foot patrol officers to a specific area, he pointed out that this would be a waste of manpower, since crime patterns indicated that other areas should have priority.

6

When two suspects threatened a false arrest suit, the sergeant advised them that they were stopped on probable cause and were free to go, though they could still make a complaint through proper channels and a follow-up investigation would be made.

5

A sergeant told a citizen who felt he had been improperly tagged that there was nothing the sergeant could do about it and the man would have to straighten it out with the judge.

4

When a sergeant requested additional equipment and his request was dropped, he enlisted the aid of his political friends to acquire the equipment.

3

A citizen involved in a car accident asked the sergeant to drive his passenger five blocks to keep an appointment. The sergeant replied, "We don't run a taxi service."

2

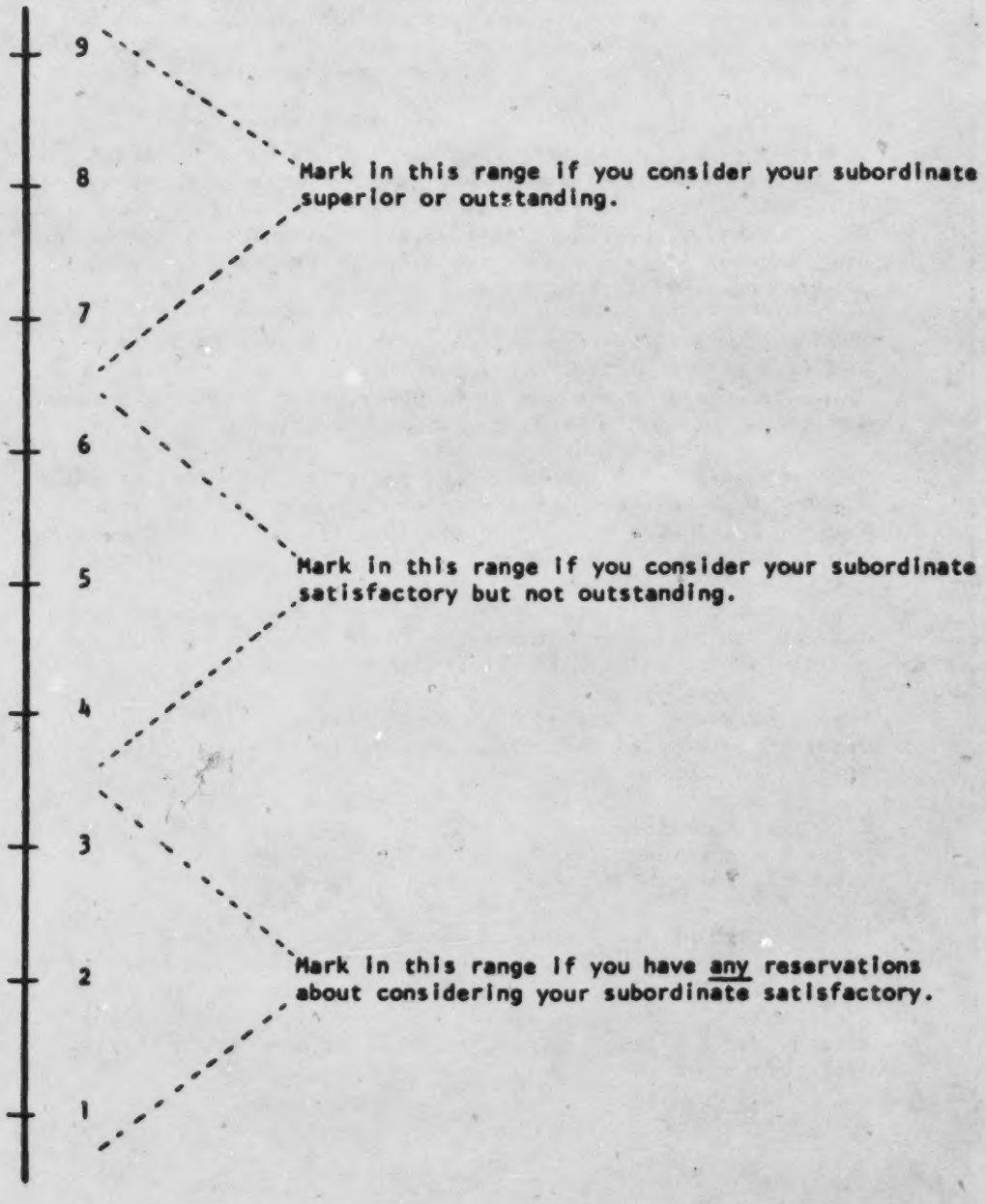
When a sergeant who was scheduled to speak at several community "open forums" failed to show up for them was asked why, he said that he was a poor public speaker and the meetings would only aggravate the bad community-police relations that already existed and, besides, the program was beyond the scope of normal police duty.

1

## OVERALL JOB PERFORMANCE

Job Category I

Consider here the overall performance of a subordinate you are rating. The following should enter into your rating: how well does he do the total job; how valuable is he to the department; how much does he contribute to the department's effectiveness.


9
8
7
6
5
4
3
2
1
Mark in this range if you consider your subordinate superior or outstanding.
Mark in this range if you consider your subordinate satisfactory but not outstanding.
Mark in this range if you have any reservations about considering your subordinate satisfactory.

JOB PERFORMANCE DESCRIPTION BOOKLET*

f o r

INTERMEDIATE COMMANDERS

* Feel free to page through this booklet before the rating session begins. The directions are more complicated than the usual rating form--however, once you get the idea, the ratings are easy to do and will go quickly.

PLEASE DO NOT START THE RATINGS UNTIL ALL DIRECTIONS HAVE BEEN GIVEN.

## JOB PERFORMANCE DESCRIPTION SCALES

### INTERMEDIATE COMMANDERS

Measuring job performance has always been an important issue for both supervisors and subordinates. Supervisors want to know how well a man is performing his job so that they can make decisions about him such as salary increases, promotions, etc. And all men, whether supervisors or subordinates, want to know how well they are doing their own job so that they can pay more attention to the things they do poorly and thereby improve.

However, measuring job performance is a difficult thing to do. Though we all speak English, we quite often have a different understanding of what some common words mean, such as good, poor, average, etc. Because we differ in what we mean by these words, we often cannot agree on whether a person's job performance is "good," "average," or "poor." Numbers don't help either, because people still have trouble agreeing on what they mean in terms of job performance. Another problem in measuring job performance is that most people find it hard to make up their minds about what the major characteristics are of what they want to measure.

Let's say you were going to paint a room brown for a friend. You show him the shade of paint you have and he says that the shade you have is "a lot less brown" than the shade he wants. So you change the combination of colors and he says that the shade you have is "not quite as brown" as the one he had in mind. You change the combination of colors that make up brown and he says, "That shade is a little less brown than I wanted." It is clear that what your friend means by "a lot less brown," "not quite as brown," "a little too brown," etc. is not the same as what you mean by "a lot less brown," "not quite as brown," etc. This is an example of the problem people have with measuring things using words that are not closely related to the thing they are measuring, in this case the color brown.

If your friend had told you that the paint was too shiny, you would have known he was talking about something other than the shade of the paint, but what if he took the shininess of the paint into account and didn't tell you? This is an example of having different ideas of what is involved in measuring brownness.

Well, you got smart and decided that you would make up a whole series of brown paint chips according to different combinations of colors and arrange them so that your friend could look at them and pick the shade he expected his room to be painted in. Of course, not every possible shade of brown was included in the colors you showed him, but the examples you had mapped out formed a scale of brown.

We have tried to develop job performance rating scales for intermediate commanders that consist of about the same thing. Each scale is a list of possible examples of behavior. The examples form an outline of what is involved in doing one aspect of the job of an intermediate commander. The examples also serve as guideposts or trail markers along the scale, marking off more and more effective performance.

We developed nine different rating scales plus an overall performance scale for the job of intermediate commander, based on information provided by intermediate commanders in a major metropolitan city. The names of these scales are:

A. Administrative Duties
B. Communications
C. Scheduling
D. Training
E. Supervision
F. Commending, Disciplining, and Assigning Efficiency Ratings
G. Field Command Situations
H. Public and Community Relations
I. Dedication, Integrity, Setting an Example
J. Overall Job Performance

Instructions: In the booklet accompanying these directions you will find ten job categories representing performance dimensions which are important for the job success of investigators. Each job category has immediately below it a definition of the category and eight performance examples gathered from policemen in a number of cities. Each performance example "anchors" the portion of the scale opposite it. That is, on the job category, "Administrative Duties," the performance example which begins, "When additional equipment was needed . . ." represents the kind of behavior or performance which should be rated at eight or nine. Thus, the performance examples should help to define further the job category and should provide "benchmarks" for defining points on the nine-point scales.

In addition to the job categories booklet containing the ten categories, their definitions, and the performance examples, we have provided you with a rating worksheet for each job category with the names of the intermediate commanders you will be rating on the left side of the worksheets. You will then use these worksheets to record the ratings you assign to your subordinates. When you are ready to begin the rating task, you should first read over the definition and the performance examples for Job Category A, "Administrative Duties." Get the content of this job category firmly in mind. Now consider the typical performance within this category of your first subordinate ratee. Compare his typical per-

formance with the performance represented by the eight "benchmarks." Assign this man the rating which best typifies his level of performance in the area of "Administrative Duties" compared with the example anchors. Record that rating opposite his name on the worksheet. Then go on to the next subordinate and rate him on the same job category. When you have finished rating all of your subordinates on the first category, turn to the next category (Job Category B, "Communications") and proceed the same way you did for the first category. Follow the same procedures for the other nine categories as well.

Things to guard against: Several sources of error can contribute to inaccuracies in your ratings. Mentioning them briefly may help you to guard against making them. Here are suggestions for overcoming them:

1) Consider each performance category separately from all the rest. An almost universal error in ratings is called HALO ERROR. It occurs when the rater gives about the same ratings to a person on all aspects of performance. Usually this error occurs because a rater has not really taken enough time to get clearly in mind what each separate category of performance refers to. Remember we are asking you to describe or rate each of your subordinates on ten categories of performance. As you consider each of the persons you are rating try to avoid getting into the habit of giving about the same rating to him on each job category. Consider each category separately from all others. Be sure to rate all your subordinates on each category before going on to the next category.

2) Consider each subordinate's performance over time and not on just one or two occasions. Another type of error occurs when a rater is influenced by just some single event or some recent occurrence. As you consider each subordinate's performance, think back over all the time you have known him and try to avoid being influenced by just one or two events. Base your ratings on all your observations of him and not just a few.

3) Avoid being misled by such things as appearance, education, family background, and other personal characteristics. Another common error in rating is called STEREOTYPE ERROR. It occurs when a rater allows himself to be influenced by other things than what the person has actually done on the job. In considering each subordinate's job performance, try to ignore everything else you may know about that person. Give your rating based strictly on what the individual has done on the job.

4) Avoid using your own definitions for the various job categories. A common reason for inaccurate ratings is because raters have different definitions of the job categories. Terms such as "Administrative Duties," "Communications," and "Supervision," etc., can have different meanings for different raters. This is why it is so very important for you to read the definitions and performance examples carefully for these job categories. Avoid any previous impressions of what these things have meant to you. Base your ratings on the definitions which are provided in the rating booklet you have received.

Remember, these ratings will be used solely for the purpose of validating the experimental battery of tests given earlier to those persons you are rating. Neither performance evaluations nor test scores will be shown to any member of the police department. Thus, you should feel free to be open and honest in completing these ratings. Thank you very much for your cooperation.

## ADMINISTRATIVE DUTIES

**Job Category A**

Preparing, reviewing, critiquing, and forwarding reports; setting up procedures for handling paper work; supervising equipment and station maintenance, requesting equipment; keeping appropriate and up-to-date records; analyzing computer statistics.

**9** When additional equipment was needed for the precinct, the Intermediate Commander asked his immediate subordinates to write a request detailing the amount of equipment needed and the reasons for the equipment. He then added his own recommendations and thoughts and forwarded the request to Headquarters with the names of all the men who worked on it.

**8** When the Intermediate Commander who was responsible for all of the station records found that the job was too much for one man to handle, he assigned record keeping duties to sergeants, based on the kinds of records the sergeants dealt with directly. He then spot checked the sergeants' record keeping.

**7** The Intermediate Commander assigned one man the additional duty of coordinating all requests for equipment. The Intermediate Commander also made this man responsible for processing, receiving, and inventorying the equipment.

**6** When the Intermediate Commander learned that all equipment had not been transferred to a new squad car when the old car was replaced, he made arrangements with the garage to have all equipment transferred.

**5** Because the station janitor did not work on weekends, the Intermediate Commander did light cleaning and sweeping to keep the station clean.

**4** The Intermediate Commander asked personnel to submit a list of divisional equipment needs, but then adopted a policy of "use the equipment already available, have it fixed if necessary, and make do."

**3** The Intermediate Commander never reviewed any reports, because he said it was just "busy work."

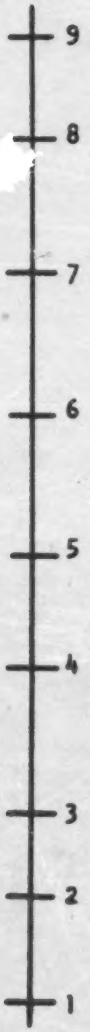
**2** An Intermediate Commander maintained a record system within his station that was unique to his station. He ignored repeated requests to upgrade his system to comply with the rest of the city, because he maintained that his system was the best.

**1**

## COMMUNICATIONS

**Job Category B**

Holding staff meetings; listening to gripes; following the chain of command; obtaining and disseminating information; providing reasons for changes to involved parties; keeping channels of communications open; exchanging information with other divisions.




9 — The Intermediate Commander held monthly staff meetings to discuss problems and needed changes. The Intermediate Commander tried to get everyone's opinion and suggestions before making a decision.

8 — When officer morale was low because rumors about plans for departmental reorganization were going around, the Intermediate Commander advised administration of the problem and requested an immediate release of the reorganization plans to dispel further rumors.

7 — Because a newly established unit was the victim of jokes and ridicule by other operation units, the Intermediate Commander visited the heads of other units and tried to clarify any misunderstandings about the new unit.

6 — When many patrolemn requested transfers after a new Intermediate Commander took over the precinct, the Intermediate Commander asked his immediate subordinates if they knew why the men wanted transfers and how he could change their minds.

5 — Although the Intermediate Commander thoroughly briefed his men on changes in procedures, he never would answer questions about the reason for change.

4 — The Intermediate Commander got a call from a newly assigned Sgt. requesting an appointment to talk about his new assignment. The Intermediate Commander would not work the Sgt. into his schedule and just told him when to report.

3 — The Intermediate Commander attended division meetings, but did not take notes or tell subordinates what was discussed.

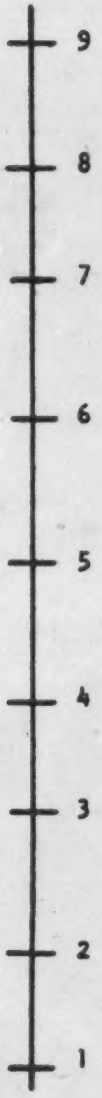
2 — A new Intermediate Commander with no supervisory experience discontinued regular staff meetings, because he said he would be making all the decisions, anyway.

1

## SCHEDULING

**Job Category C**

Matching manpower assignments with crime trends; developing crime pattern information; acting as a final arbiter in scheduling problems such as assigning days off, vacations etc.; being flexible when taking scheduling problems into account; assigning details equitably; obtaining replacements to fill manpower shortages.




9 — Because there was a high incidence of dwelling burglaries in his precinct, the Intermediate Commander used computer information on times, areas, M.O.'s etc., so he could detail surveillance men to saturate the areas at critical times.

8 — With the increasing amount of walk-in traffic to the division, the Intermediate Commander assigned one man, on a rotating basis, to work inside and handle citizen problems.

7 — The Intermediate Commander obtained a volunteer from the power shift to take a man's assignment for two weeks, because the officer's vacation had been rescheduled and he had already made reservations.

6 — When an officer asked the Intermediate Commander for a day off on short notice to attend his child's christening, the Intermediate Commander refused his request because another officer had already asked for that day off in advance and he had a minimal crew.

5 — When a patrolman asked an Intermediate Commander for a change in his beat, the Intermediate Commander said that he couldn't change the man's assignment at the time because six other patrolmen would be affected by a change.

4 — An Intermediate Commander approved an officer's request for days off for a vacation without consulting his immediate subordinate about the manpower resources for that period.

3 — The Intermediate Commander responsible for making up monthly work sheets assigned plenty of men for the days he would be working and a lighter crew for those days when he wouldn't be, making it difficult for the Sgt. to get the job done when he was working alone.

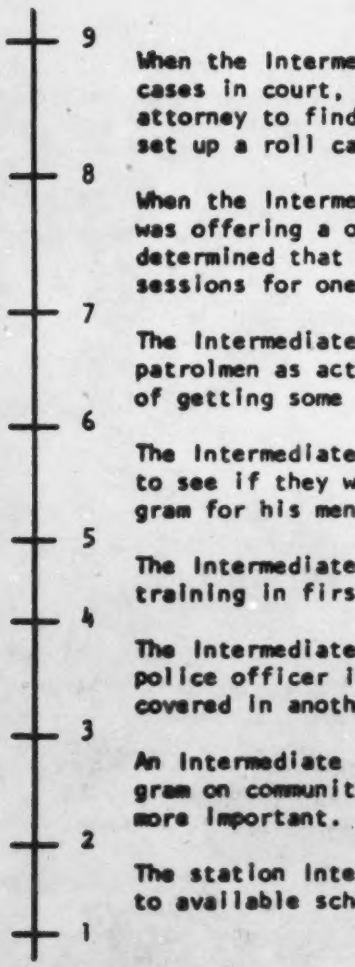
2 — The Intermediate Commander frequently gave his men extra Friday and Saturday nights off, even though these were the heavy crime nights, because he wanted to be a nice guy.

1

## TRAINING

**Job Category D**

Developing and instituting training programs; determining training needs; being aware of the training opportunities; evaluating the effect of training programs; doing on-the-job training by delegating authority; giving men experience on different jobs as a part of training.



9 When the Intermediate Commander's station or unit began losing cases in court, he contacted the judge and the prosecuting attorney to find out why. He then used their suggestions to set up a roll call training program for his men.

8 When the Intermediate Commander was informed that the county coroner was offering a one-hour training session on handling D.O.A.'s, he determined that there was a need for the training and set up the sessions for one hour before roll call.

7 The Intermediate Commander set up a rotating policy of assigning patrolmen as acting Sgt. so that every qualified man had a chance of getting some training in supervising.

6 The Intermediate Commander closely reviewed all training materials to see if they were worthwhile before scheduling the training program for his men.

5 The Intermediate Commander did not ask if anyone needed or wanted training in first aid, but just set up some training in it.

4 The Intermediate Commander instituted a training program for every police officer in the station which covered material that had been covered in another program three months earlier.

3 An Intermediate Commander refused to schedule a short training program on community relations because he felt that other things were more important.

2 The station Intermediate Commander would not send any of his officers to available schools, because he said that he didn't believe in them.

1

## SUPERVISION

**Job Category E**

Observing subordinates' performance; specifying proper procedures for men to follow; following up to see if orders were carried out; delegating authority to others; providing feedback on day-to-day performance to subordinates.

**9** When an intermediate commander observed an officer make an illegal arrest, he immediately took the officer aside and told him how to correct his mistake.

**8** The Sgt. had been pushing a group of poor performers hard and they complained to the intermediate commander. The intermediate commander told them that if he were the Sgt. he would be doing the same thing, and that the pressure would continue until they shaped up.

**7** Because the intermediate commander was concerned with the quality of arrests, he issued a policy stating that the sergeants must be present at the scene of felony arrests or must immediately review all felony arrests.

**6** The intermediate commander continuously checked every case assigned to his division to be sure that it was handled correctly, and frequently called his men to be sure they had completed their responsibilities.

**5** When a Sgt. requested that a patrolman who wasn't performing effectively be taken off squad duty and assigned to a beat, the intermediate commander told the Sgt. to do as he pleased with the patrolman.

**4** The intermediate commander did not permit his officers to change their assigned schedules without his consent, and he did not delegate any authority to his subordinates.

**3** An intermediate commander always second-guessed his men and made decisions for them. That is, he always asked them if they had done this or that, and then told them what they should do and how to go about it.

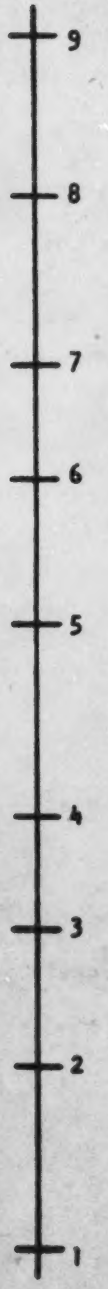
**2** The intermediate commander allowed a Sgt. to harass and berate some newly appointed patrolmen, because he thought getting chewed out and razzed would be good for them, even if they didn't really deserve it.

**1**

## COMMENDING, DISCIPLINING, AND ASSIGNING EFFICIENCY RATINGS

**Job Category F**

Obtaining factual information before acting on commendations or disciplinary actions; disciplining men in private; making commendations public; fitting the disciplinary action to the incident; facing up to disciplinary problems; discussing ratings with immediate supervisors; reviewing ratings with men; being unbiased when making ratings.



**9** The Intermediate Commander reviewed all arrest reports and conferred with his immediate subordinates on instances of outstanding performance, recommending supervisors place selected men in for commendations. He followed up these incidents to make sure the commendations request was submitted.

**8** When a patrolman who had a drinking problem was found drinking on duty, the Intermediate Commander referred him to higher authorities for discipline, but requested that he not be transferred. The Intermediate Commander also assisted the man in joining an AA group.

**7** The Intermediate Commander called his immediate subordinates into his office to discuss discrepancies and differences between their efficiency ratings.

**6** When the Intermediate Commander discovered that a squad had stopped to talk to a citizen on the way to an "assault in progress" (and the assailant had been apprehended by the backup squad), he recommended to his superior that disciplinary action be taken.

**5** When officers apprehended a suspect in a remote area with the help of a police dog, a Sgt. recommended that the men be given citations. The Intermediate Commander, however, upon reviewing the facts, decided that only the dog's action was outstanding and the officers were only performing routine functions, so he decided against a citation.

**4** The Intermediate Commander gave high ratings only to officers who did something spectacular or heroic. The steady performers who did their jobs just got average ratings.

**3** When the Intermediate Commander received a complaint from a community leader that one of his patrolmen had struck a black youth without provocation, he called the officer in and severely reprimanded him without asking for his side of the story.

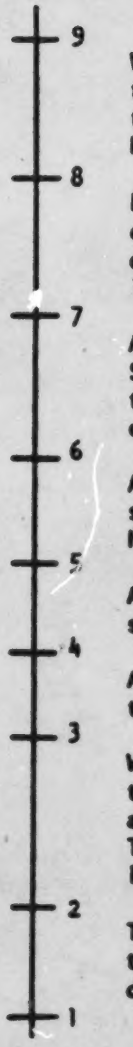
**2** An Intermediate Commander received a written report from a field Sgt. that an officer was drinking on duty, but he tore up the report. A week later when the inspector observed the problem and told the Intermediate Commander, the Intermediate Commander chewed out the Sgt. for not stopping the problem.

**1**

## FIELD COMMAND SITUATIONS

Job Category G

Making decisions about courses of action; coordinating personnel at a crime scene, civil disturbance, disaster, etc.; deploying men to control the scene; exercising caution in dangerous situations; supervising a field situation until resolved or properly relieved.



9

When a serious disturbance arose which required cars from all over the city to assist, the Intermediate Commander had a Sgt. take over the station and went to the scene to set up a temporary command post. He directed activities until he was relieved by an Inspector.

8

Because a Sgt. was not available for on-scene supervision at a civil disturbance, the Intermediate Commander immediately responded and took command of the scene with his own officers and without having to call in other supervisors or units.

7

An Intermediate Commander at a fatal MVA took charge by ordering a Sgt. to handle rescue operations and he himself posted traffic control officers and had firemen rope off a pedestrian area to keep the crowd under control.

6

An Intermediate Commander talked with investigators at the scene of a shooting and took responsibility for protecting the scene. He instructed his men, deployed them, and then left.

5

At the scene of a bomb threat, the Intermediate Commander did not send his men to evacuate adjoining buildings.

4

An Intermediate Commander told a Sgt. to take command of the men at the scene of a factory explosion and then left.

3

When two men fled from a robbery and entered an unlocked apartment, the Intermediate Commander ordered tear gas to be fired into the apartment, but he didn't ask the tenant which windows were hers. Thus, tear gas was fired into the adjacent apartments as well, causing expensive clean-up costs.

2

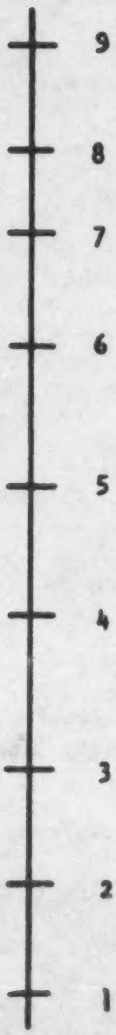
The Intermediate Commander entered the scene of a homicide and touched numerous articles that might have been valuable as evidence had he left them alone.

1

## PUBLIC AND COMMUNITY RELATIONS

Job Category H

Dealing courteously and constructively with community problems; working with community groups; dealing with the public both in the station and out on the street; providing a good public image.



9

When an Intermediate Commander took over a precinct, he went to all of the schools, introduced himself to the principals, and told them to feel free to call him if he could be of service.

8

When a well-known community leader died, the Intermediate Commander sent flowers to the funeral parlor and then paid his respects.

7

An Intermediate Commander told a group of black community leaders that he didn't care if a lawbreaker was black, red, green, or blue--he still would be treated like any other lawbreaker.

6

The Intermediate Commander called a citizen who had registered a complaint against one of his men and explained to him the officer's version of the story to show him something was being done about the incident.

5

When a large PTA group called the Intermediate Commander and asked for him to speak to them on youth and drugs, he referred them to the Juvenile Division without following up to see that they got a speaker.

4

A group of businessmen requested an appointment with the Intermediate Commander to discuss problems in the area. The Intermediate Commander set up the appointment and then had his immediate subordinate, who was also familiar with the area's problems, keep the appointment.

3

The Intermediate Commander kept a community leader who had a complaint about a recent incident waiting for 45 minutes while he finished his lunch.

2

The Intermediate Commander refused to talk to a group of concerned citizens about community problems, stating this was not part of his job.

1

## DEDICATION, INTEGRITY, SETTING AN EXAMPLE

Job Category I

Serving as a model for subordinates; abiding by all department rules and regulations; accepting rather than avoiding one's job responsibilities; avoiding acts that might be construed as compromising one's integrity.

9

An Intermediate Commander postponed his vacation because there was an outbreak of bombing and bomb threats.

8

The Intermediate Commander took night school courses to improve his managerial skills.

7

The Intermediate Commander made periodic checks of the station at various shifts to make closer contact with his field personnel to see if there were any problems he could help solve.

6

The Intermediate Commander gave a full eight hours to the job, but he rarely became involved in anything beyond his daily duties.

5

The Intermediate Commander accepted a case of liquor for the station party from a liquor store in the precinct.

4

An Intermediate Commander came to work at 10 a.m. wearing uniform trousers and a t-shirt. He called for a squad to pick up his shirts at the laundry and bring them to the station so he could get dressed for the job.

3

An Intermediate Commander conducting inspection berated officers for their sloppy appearance, even though he was wearing an unpressed, unclean uniform, unauthorized wing-tip shoes, and other nonregulation items, and had a cigar in his mouth.

2

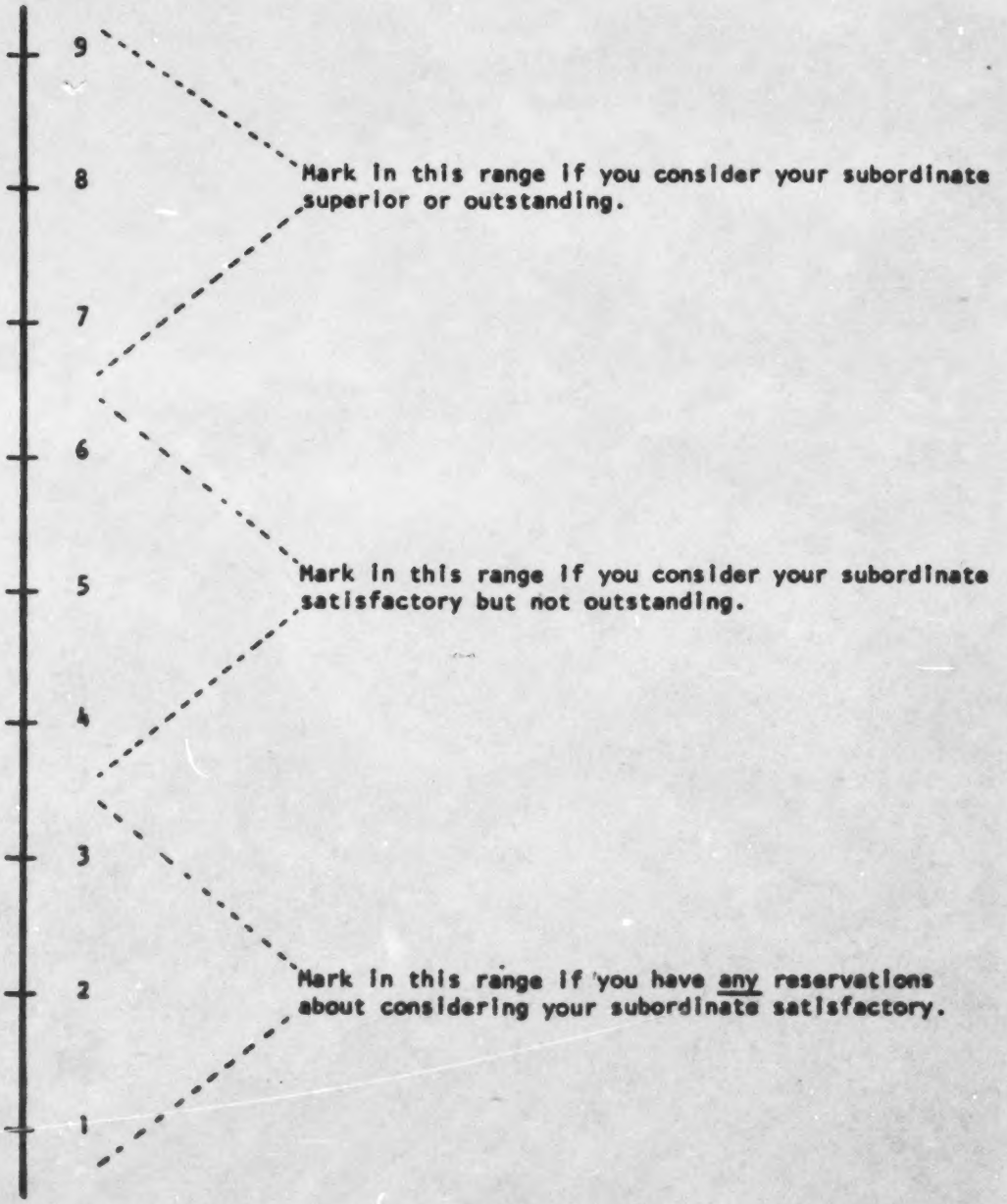
An officer arrested and brought in two black suspects and explained the charges to the Intermediate Commander. The Intermediate Commander realized the charges were too vague, but he booked them to avoid being called a "nigger lover."

1

## OVERALL JOB PERFORMANCE

Job Category J

Consider here the overall performance of a subordinate you are rating. The following should enter into your rating: how well does he do the total job; how valuable is he to the department; how much does he contribute to the department's effectiveness.


9
8
7
6
5
4
3
2
1
Mark in this range if you consider your subordinate superior or outstanding.
Mark in this range if you consider your subordinate satisfactory but not outstanding.
Mark in this range if you have any reservations about considering your subordinate satisfactory.

# APPENDIX B

# TABLES SHOWING INTERRATER RELIABILITY ESTIMATES FOR JOB PERFORMANCE RATINGS SEPARATELY, BY CITY, FOR DETECTIVES, SERGEANTS, AND MIDDLE COMMAND PERSONNEL

TABLE 1

*Interrater Reliability Estimates* for Job Performance Ratings for Cincinnati Detectives*
(N = 65)

| Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=19) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|
| A Investigating a Crime | 62 | 78 | 67 |
| B Arrest, Search, Seizure | 61 | 75 | 66 |
| C Interrogation | 71 | 70 | 76 |
| D Investigating a Case | 52 | 68 | 57 |
| E Developing Informants | 72 | 76 | 76 |
| F Report Writing | 45 | 52 | 51 |
| G Appearing in Court | 0 | 05 | 0 |
| H Public Relations | 48 | 58 | 54 |
| I Juveniles | 13 | 36 | 17 |
| J Cooperation | 58 | 62 | 63 |
| K Dedication | 48 | 72 | 55 |
| L Integrity | 02 | 24 | 02 |
| M Overall Rating | 60 | 73 | 66 |
| Sum of Ratings | 68 | 78 | 73 |
| Mean Number of Raters (K) | | 1.55 | |

*Decimals omitted.

TABLE 2

*Interrater Reliability Estimates* for Job Performance Ratings for Des Moines Detectives*
(N = 33)

| Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=21) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|
| A Investigating a Crime | 30 | −26 | 32 |
| B Arrest, Search, Seizure | −51 | −1.35 | 0 |
| C Interrogation | 40 | 45 | 42 |
| D Investigating a Case | 64 | 57 | 67 |
| E Developing Informants | 59 | 54 | 62 |
| F Report Writing | 18 | 05 | 20 |
| G Appearing in Court | 43 | 27 | 46 |
| H Public Relations | 29 | −12 | 31 |
| I Juveniles | 28 | 15 | 30 |
| J Cooperation | 51 | 11 | 53 |
| K Dedication | 45 | 47 | 48 |
| L Integrity | −85 | −46 | 0 |
| M Overall Rating | 65 | 49 | 68 |
| Sum of Ratings | 48 | 30 | 51 |
| Mean Number of Raters (K) | 1.80 | | |

*Decimals omitted.

TABLE 3

*Interrater Reliability Estimates* for Job Performance Ratings for Miami (Dade Co.) Detectives*
(N = 95)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=52) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Investigating a Crime | 48 | 46 | 54 |
| B | Arrest, Search, Seizure | 59 | 70 | 64 |
| C | Interrogation | 35 | 45 | 40 |
| D | Investigating a Case | 51 | 54 | 57 |
| E | Developing Informants | −13 | −27 | 0 |
| F | Report Writing | −33 | −38 | 0 |
| G | Appearing in Court | 14 | −02 | 18 |
| H | Public Relations | −23 | −06 | 0 |
| I | Juveniles | −39 | −34 | 0 |
| J | Cooperation | 28 | 19 | 33 |
| K | Dedication | −11 | 0 | 0 |
| L | Integrity | −39 | −1.26 | 0 |
| M | Overall Rating | 56 | 56 | 62 |
| | Sum of Ratings | 24 | 28 | 29 |
| | Mean Number of Raters (K) | 1.60 | 1 | |

*Decimals omitted.

TABLE 4

*Interrater Reliability Estimates* for Job Performance Ratings for Minneapolis Detectives*
(N = 55)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=31) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Investigating a Crime | 80 | 80 | 80 |
| B | Arrest, Search, Seizure | 65 | 73 | 65 |
| C | Interrogation | 56 | 65 | 55 |
| D | Investigating a Case | 70 | 66 | 69 |
| E | Developing Informants | 53 | 41 | 52 |
| F | Report Writing | 57 | 66 | 56 |
| G | Appearing in Court | 25 | 06 | 25 |
| H | Public Relations | 64 | 69 | 63 |
| I | Juveniles | 28 | 12 | 39 |
| J | Cooperation | 55 | 49 | 54 |
| K | Dedication | 60 | 58 | 59 |
| L | Integrity | 50 | 56 | 49 |
| M | Overall Rating | 75 | 69 | 74 |
| | Sum of Ratings | 74 | 74 | 73 |
| | Mean Number of Raters (K) | 2.00 | | |

*Decimals omitted.

TABLE 5

*Interrater Reliability Estimates* for Job Performance Ratings for Portland Detectives*
(N= 58)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=16) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Investigating a Crime | 67 | 70 | 74 |
| B | Arrest, Search, Seizure | 23 | 06 | 29 |
| C | Interrogation | 15 | −09 | 20 |
| D | Investigating a Case | 53 | −08 | 61 |
| E | Developing Informants | 23 | 48 | 29 |
| F | Report Writing | 44 | 06 | 52 |
| G | Appearing in Court | 15 | −32 | 21 |
| H | Public Relations | 25 | −14 | 32 |
| I | Juveniles | 13 | 08 | 18 |
| J | Cooperation | 27 | −45 | 34 |
| K | Dedication | 63 | 43 | 70 |
| L | Integrity | 57 | 55 | 64 |
| M | Overall Rating | 58 | 58 | 66 |
| | Sum of Ratings | 50 | 40 | 58 |
| | Mean Number of Raters (K) | 1.45 | | |

*Decimals omitted.

TABLE 6

*Interrater Reliability Estimates* for Job Performance Ratings for Washington, D. C. Detectives*
(N = 90)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=30) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Investigating a Crime | 66 | 71 | 72 |
| B | Arrest, Search, Seizure | 27 | 35 | 33 |
| C | Interrogation | 44 | 35 | 52 |
| D | Investigating a Case | 46 | 52 | 54 |
| E | Developing Informants | 14 | −39 | 19 |
| F | Report Writing | 18 | −25 | 23 |
| G | Appearing in Court | 55 | 31 | 62 |
| H | Public Relations | −03 | −34 | 0 |
| I | Juveniles | 05 | −07 | 08 |
| J | Cooperation | 28 | −06 | 34 |
| K | Dedication | 25 | −76 | 31 |
| L | Integrity | 12 | 21 | 15 |
| M | Overall Rating | 30 | 38 | 36 |
| | Sum of Ratings | 42 | 28 | 49 |
| | Mean Number of Raters (K) | 1.50 | | |

*Decimals omitted.

TABLE 7

*Interrater Reliability Estimates* for Job Performance Ratings for Cincinnati Sergeants*
(N = 74)

| | Dimension | Overall Estimate | Estimates Based on Raters with Two Ratings Only (N=44) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | 30 | 24 | 34 |
| B | Coordination and Deployment | 48 | 42 | 54 |
| C | Supervision | 42 | 38 | 47 |
| D | Administration and Inspection | 04 | 09 | 05 |
| E | Decision Making and Initiative | 30 | 29 | 35 |
| F | Training and Planning | 49 | 55 | 55 |
| G | Dedication | 68 | 71 | 75 |
| H | Public Contact | 52 | 56 | 58 |
| I | Overall Rating | 64 | 65 | 69 |
| | Sum of Ratings | 52 | 53 | 58 |
| | Mean Number of Raters (K) | 1.60 | | |

*Decimals omitted.

TABLE 8

*Interrater Reliability Estimates* for Job Performance Ratings for Des Moines Sergeants*
(N = 46)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=30) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | 34 | 45 | 39 |
| B | Coordination and Deployment | 55 | 60 | 60 |
| C | Supervision | 64 | 71 | 68 |
| D | Administration and Inspection | 62 | 74 | 67 |
| E | Decision Making and Initiative | 28 | 43 | 33 |
| F | Training and Planning | 49 | 53 | 54 |
| G | Dedication | 55 | 65 | 60 |
| H | Public Contact | 63 | 65 | 67 |
| I | Overall Rating | 70 | 75 | 74 |
| | Sum of Ratings | 72 | 78 | 75 |
| | Mean Number of Raters (K) | 1.65 | | |

*Decimals omitted.

TABLE 9

*Interrater Reliability Estimates* for Job Performance Ratings for Miami (Dade Co.) Sergeants*
(N = 51)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=32) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | 53 | 60 | 58 |
| B | Coordination and Deployment | 61 | 56 | 50 |
| C | Supervision | 67 | 70 | 71 |
| D | Administration and Inspection | 46 | 27 | 51 |
| E | Decision Making and Initiative | 43 | 45 | 48 |
| F | Training and Planning | 56 | 61 | 61 |
| G | Dedication | 36 | 52 | 40 |
| H | Public Contact | 43 | 40 | 47 |
| I | Overall Rating | 59 | 61 | 63 |
| | Sum of Ratings | 69 | 68 | 72 |
| | Mean Number of Raters (K) | 1.65 | | |

*Decimals omitted.

TABLE 10

*Interrater Reliability Estimates* for Job Performance Ratings for Minneapolis Sergeants*
(N = 40)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=13) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | 55 | 54 | 44 |
| B | Coordination and Deployment | 45 | 47 | 34 |
| C | Supervision | 37 | 34 | 28 |
| D | Administration and Inspection | 37 | 03 | 27 |
| E | Decision Making and Initiative | 47 | 33 | 37 |
| F | Training and Planning | 63 | 85 | 53 |
| G | Dedication | 40 | −22 | 30 |
| H | Public Contact | 50 | 04 | 40 |
| I | Overall Rating | 62 | 63 | 52 |
| | Sum of Ratings | 61 | 47 | 50 |
| | Mean Number of Raters (K) | 3.10 | | |

*Decimals omitted.

TABLE 11

*Interrater Reliability Estimates* for Job Performance Ratings for Portland Sergeants*
(N = 16)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=7) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | −16 | 02 | 0 |
| B | Coordination and Deployment | −47 | −08 | 0 |
| C | Supervision | 52 | 63 | 57 |
| D | Administration and Inspection | 38 | 08 | 42 |
| E | Decision Making and Initiative | 28 | −80 | 32 |
| F | Training and Planning | −08 | 56 | 0 |
| G | Dedication | −07 | −05 | 0 |
| H | Public Contact | −57 | −29 | 0 |
| I | Overall Rating | −13 | 63 | 0 |
| | Sum of Ratings | 10 | 55 | 12 |
| | Mean Number of Raters (K) | 1.67 | | |

*Decimals omitted.

TABLE 12

*Interrater Reliability Estimates* for Job Performance Ratings for San Diego Sergeants*
(N = 52)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=26) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | 29 | 01 | 34 |
| B | Coordination and Deployment | 11 | −10 | 13 |
| C | Supervision | 46 | 29 | 33 |
| D | Administration and Inspection | 49 | 48 | 55 |
| E | Decision Making and Initiative | 40 | −48 | 46 |
| F | Training and Planning | 28 | 04 | 33 |
| G | Dedication | 36 | 29 | 43 |
| H | Public Contact | 06 | 39 | 08 |
| I | Overall Rating | 24 | 01 | 30 |
| | Sum of Ratings | 34 | 10 | 40 |
| | Mean Number of Raters (K) | 1.50 | | |

*Decimals

TABLE 13

*Interrater Reliability Estimates* for Job Performance Ratings for Washington, D. C. Sergeants*
(N = 88)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=24) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Concern for Subordinates | 30 | 43 | 39 |
| B | Coordination and Deployment | 32 | 20 | 41 |
| C | Supervision | 46 | 56 | 55 |
| D | Administration and Inspection | 51 | 54 | 64 |
| E | Decision Making and Initiative | 10 | 44 | 14 |
| F | Training and Planning | 49 | 45 | 58 |
| G | Dedication | 12 | 37 | 17 |
| H | Public Contact | 20 | 41 | 27 |
| I | Overall Rating | 70 | 80 | 78 |
| | Sum of Ratings | 63 | 71 | 71 |
| | Mean Number of Raters (K) | 1.35 | | |

*Decimals omitted.

TABLE 14

*Interrater Reliability Estimates* for Job Performance Ratings for Des Moines Middle Level Officers*
(N = 23)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=8) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Administrative | 60 | 68 | 69 |
| B | Communications | 09 | 08 | 13 |
| C | Scheduling | 64 | 67 | 72 |
| D | Training | 59 | 68 | 69 |
| E | Supervision | 0 | −1.90 | 0 |
| F | Subordinate Evaluation | 26 | −71 | 34 |
| G | Field Command | −33 | −20 | 0 |
| H | Public Relations | 39 | 58 | 49 |
| I | Dedication | 34 | 23 | 44 |
| J | Overall Performance | 45 | 28 | 55 |
| | Sum of Ratings | 59 | 43 | 68 |
| | Mean Number of Raters (K) | 1.34 | | |

*Decimals omitted.

TABLE 15

*Interrater Reliability Estimates* for Job Performance Ratings for Miami (Dade Co.) Middle Level Officers*
(N = 31)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=10) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Administrative | 47 | 27 | 57 |
| B | Communications | 43 | 27 | 53 |
| C | Scheduling | 09 | −40 | 13 |
| D | Training | 38 | −1.00 | 49 |
| E | Supervision | 37 | 41 | 48 |
| F | Subordinate Evaluation | 85 | 89 | 89 |
| G | Field Command | 0 | −38 | 0 |
| H | Public Relations | −31 | −2.10 | 0 |
| I | Dedication | 51 | 67 | 61 |
| J | Overall Performance | 24 | 38 | 33 |
| | Sum of Ratings | 46 | 39 | 56 |
| | Mean Number of Raters (K) | 1.32 | | |

*Decimals omitted.

TABLE 16

*Interrater Reliability Estimates* for Job Performance Ratings for Minneapolis Middle Level Officers*
(N = 28)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=12) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Administrative | 63 | 65 | 71 |
| B | Communications | 77 | 87 | 83 |
| C | Scheduling | 28 | −16 | 35 |
| D | Training | 64 | 77 | 72 |
| E | Supervision | 01 | −39 | 01 |
| F | Subordinate Evaluation | 66 | 67 | 73 |
| G | Field Command | 24 | −97 | 30 |
| H | Public Relations | 10 | 40 | 14 |
| I | Dedication | 53 | 23 | 61 |
| J | Overall Performance | 68 | 71 | 75 |
| | Sum of Ratings | 55 | 62 | 63 |
| | Mean Number of Raters (K) | 1.42 | | |

*Decimals omitted.

TABLE 17

*Interrater Reliability Estimates* for Job Performance Ratings for Portland Middle Level Officers*
(N = 21)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=12) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Administrative | 64 | 52 | 65 |
| B | Communications | 65 | 19 | 66 |
| C | Scheduling | 22 | 62 | 23 |
| D | Training | 37 | 55 | 38 |
| E | Supervision | 78 | 75 | 79 |
| F | Subordinate Evaluation | 40 | 14 | 42 |
| G | Field Command | 59 | 42 | 61 |
| H | Public Relations | 02 | −49 | 02 |
| I | Dedication | 58 | 64 | 60 |
| J | Overall Performance | 69 | 46 | 70 |
| | Sum of Ratings | 67 | 48 | 68 |
| | Mean Number of Raters (K) | 1.89 | | |

*Decimals omitted.

TABLE 18

*Interrater Reliability Estimates* for Job Performance Ratings for Washington, D. C. Middle Level Officers*
(N = 47)

| | Dimension | Overall Estimate | Estimate Based on Raters with Two Ratings Only (N=16) | Estimate Based on Spearman-Brown Correction to Two Raters (K=2) |
|---|---|---|---|---|
| A | Administrative | 50 | 75 | 56 |
| B | Communications | 38 | 48 | 43 |
| C | Scheduling | 38 | 22 | 44 |
| D | Training | 31 | −35 | 37 |
| E | Supervision | 55 | 68 | 61 |
| F | Subordinate Evaluation | 46 | 76 | 53 |
| G | Field Command | 46 | −81 | 53 |
| H | Public Relations | 0 | 20 | 0 |
| I | Dedication | 12 | 18 | 15 |
| J | Overall Performance | 59 | 52 | 66 |
| | Sum of Ratings | 58 | 66 | 64 |
| | Mean Number of Raters (K) | 1.56 | | |

*Decimals omitted.

# APPENDIX C

## SAMPLE SIZES, MEANS, AND STANDARD DEVIATIONS OF JOB PERFORMANCE DIMENSION RATINGS; AND CORRELATION MATRICES BETWEEN PERFORMANCE RATINGS FOR EACH OF FOUR POLICE JOB AREAS

TABLE 1

*Sample Sizes, Means, and Standard Deviations of Performance Ratings for Patrol Officers With Both Test and Rating Information*

| Dimension | N | Mean | Standard Deviation |
|---|---|---|---|
| A Crime Prevention | 618 | 6.10 | 1.36 |
| B Use of Force | 626 | 6.06 | 1.26 |
| C Traffic Control | 591 | 6.36 | 1.44 |
| D Public Safety | 465 | 6.53 | 1.18 |
| E Investigation | 363 | 6.84 | 1.28 |
| F Report Writing | 702 | 6.28 | 1.30 |
| G Ethical Conduct | 593 | 7.12 | 1.18 |
| H Dealing with the Public | 542 | 6.38 | 1.17 |
| I Domestic Disputes | 360 | 6.49 | 1.19 |
| J Commitment | 522 | 6.66 | 1.57 |
| K Teamwork | 520 | 7.02 | 1.34 |
| L Overall Rating | 700 | 6.31 | 1.37 |
| M Mean of All Ratings | 705 | 6.47 | 1.02 |

TABLE 2

*Correlation Coefficients* (Based on Pair-Wise Deletion) Between Job Performance Rating Dimensions for Patrol Officers With Both Test and Rating Information*

| Dimension | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A Crime Prevention | .. | | | | | | | | | | | | |
| B Use of Force | 32 | .. | | | | | | | | | | | |
| C Traffic Control | 61 | 34 | .. | | | | | | | | | | |
| D Public Safety | 60 | 25 | 61 | .. | | | | | | | | | |
| E Investigation | 76 | 28 | 60 | 69 | .. | | | | | | | | |
| F Report Writing | 56 | 31 | 52 | 62 | 71 | .. | | | | | | | |
| G Ethical Conduct | 48 | 28 | 49 | 50 | 47 | 44 | .. | | | | | | |
| H Dealing with the Public | 50 | 56 | 56 | 50 | 52 | 50 | 50 | .. | | | | | |
| I Domestic Disputes | 52 | 62 | 57 | 51 | 53 | 55 | 39 | 73 | .. | | | | |
| J Commitment | 63 | 23 | 58 | 68 | 70 | 58 | 55 | 54 | 54 | .. | | | |
| K Teamwork | 49 | 18 | 54 | 66 | 67 | 51 | 40 | 56 | 57 | 60 | .. | | |
| L Overall Rating | 73 | 36 | 70 | 65 | 77 | 63 | 58 | 59 | 60 | 75 | 62 | .. | |
| M Mean of All Ratings | 79 | 52 | 79 | 75 | 83 | 75 | 69 | 76 | 75 | 82 | 73 | 88 | .. |

*Decimals omitted.

163

TABLE 3

*Sample Sizes, Means, and Standard Deviations of Performance Ratings for Detectives With Both Test and Rating Information*

| | Dimension | N | Mean | Standard Deviation |
|---|---|---|---|---|
| A | Investigating a Crime | 377 | 6.45 | 1.56 |
| B | Arrest, Search, Seizure | 334 | 6.59 | 1.48 |
| C | Interrogation | 361 | 6.39 | 1.44 |
| D | Investigating a Case | 411 | 6.49 | 1.56 |
| E | Developing Informants | 209 | 6.18 | 1.89 |
| F | Report Writing | 233 | 6.54 | 1.47 |
| G | Appearing in Court | 104 | 6.84 | 1.29 |
| H | Public Relations | 294 | 6.62 | 1.46 |
| I | Juveniles | 178 | 6.62 | 1.53 |
| J | Cooperation | 219 | 6.63 | 1.56 |
| K | Dedication | 205 | 6.32 | 1.70 |
| L | Integrity | 214 | 7.63 | 1.46 |
| M | Overall Rating | 415 | 6.51 | 1.45 |
| N | Mean of All Ratings | 415 | 6.59 | 1.15 |

TABLE 4

*Correlation Coefficients* (Based on Pair-Wise Deletion) Between Job Performance Rating Dimensions for Detectives With Both Test and Rating Information*

| | Dimension | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | Investigating a Crime | .. | | | | | | | | | | | | | |
| B | Arrest, Search, Seizure | 72 | .. | | | | | | | | | | | | |
| C | Interrogation | 63 | 61 | .. | | | | | | | | | | | |
| D | Investigating a Case | 68 | 62 | 61 | .. | | | | | | | | | | |
| E | Developing Informants | 57 | 57 | 58 | 62 | .. | | | | | | | | | |
| F | Report Writing | 53 | 57 | 36 | 55 | 39 | .. | | | | | | | | |
| G | Appearing in Court | 59 | 66 | 63 | 69 | 61 | 68 | .. | | | | | | | |
| H | Public Relations | 38 | 43 | 53 | 40 | 38 | 41 | 61 | .. | | | | | | |
| I | Juveniles | 35 | 34 | 44 | 43 | 48 | 27 | 40 | 50 | .. | | | | | |
| J | Cooperation | 51 | 57 | 54 | 58 | 43 | 51 | 58 | 59 | 45 | .. | | | | |
| K | Dedication | 73 | 68 | 61 | 75 | 60 | 60 | 65 | 50 | 50 | 58 | .. | | | |
| L | Integrity | 39 | 35 | 36 | 32 | 4? | 44 | 72 | 34 | 55 | 51 | 47 | .. | | |
| M | Overall Rating | 72 | 69 | 65 | 74 | 62 | 57 | 73 | 53 | 41 | 64 | 78 | 49 | .. | |
| N | Mean of All Ratings | 81 | 80 | 78 | 83 | 75 | 68 | 83 | 65 | 60 | 75 | 85 | 58 | 88 | .. |

*Decimals omitted.

TABLE 5

*Sample Sizes, Means, and Standard Deviations of Performance Ratings for Sergeants With Both Test and Rating Information*

| | Dimension | N | Mean | Standard Deviation |
|---|---|---|---|---|
| A | Concern for Subordinates | 350 | 7.05 | 1.43 |
| B | Coordination and Deployment | 298 | 6.39 | 1.71 |
| C | Supervision | 359 | 6.85 | 1.59 |
| D | Administration and Inspection | 284 | 6.43 | 1.76 |
| E | Decision Making Initiative | 349 | 6.72 | 1.39 |
| F | Training and Planning | 363 | 6.32 | 1.68 |
| G | Dedication | 352 | 7.14 | 1.41 |
| H | Public Contact | 352 | 6.76 | 1.42 |
| I | Overall Rating | 362 | 6.45 | 1.46 |
| J | Mean of All Ratings | 365 | 6.71 | 1.17 |

TABLE 6

*Correlation Coefficients* (Based on Pair-Wise Deletion) Between Job Performance Rating Dimensions for Sergeants With Both Test and Rating Information*

| | Dimension | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | Concern for Subordinates | .. | | | | | | | | | |
| B | Coordination and Deployment | 49 | .. | | | | | | | | |
| C | Supervision | 48 | 64 | .. | | | | | | | |
| D | Administration and Inspection | 36 | 57 | 59 | .. | | | | | | |
| E | Decision Making Initiative | 48 | 52 | 58 | 49 | .. | | | | | |
| F | Training and Planning | 51 | 55 | 53 | 58 | 46 | .. | | | | |
| G | Dedication | 47 | 48 | 45 | 48 | 52 | 54 | .. | | | |
| H | Public Contact | 50 | 45 | 47 | 46 | 41 | 56 | 55 | .. | | |
| I | Overall Rating | 60 | 66 | 64 | 68 | 66 | 70 | 69 | 59 | .. | |
| J | Mean of All Ratings | 70 | 79 | 78 | 77 | 74 | 79 | 75 | 72 | 89 | .. |

*Decimals omitted.

TABLE 7

*Sample Sizes, Means, and Standard Deviations of Performance Ratings for Middle Level Officers With Both Test and Rating Information*

| Dimension | N | Mean | Standard Deviation |
|---|---|---|---|
| A Administrative | 204 | 6.62 | 1.50 |
| B Communications | 186 | 6.62 | 1.55 |
| C Scheduling | 172 | 6.59 | 1.71 |
| D Training | 204 | 6.50 | 1.64 |
| E Supervision | 158 | 6.85 | 1.45 |
| F Subordinate Evaluation | 186 | 6.83 | 1.62 |
| G Field Command | 150 | 7.10 | 1.59 |
| H Public Relations | ... | Deleted | .... |
| I Dedication | 157 | 6.79 | 1.45 |
| J Overall Performance | 204 | 6.32 | 1.62 |
| K Mean of All Ratings | 206 | 6.65 | 1.21 |

TABLE 8

*Correlation Coefficients* (Based on Pair-Wise Deletion) Between Job Performance Rating Dimensions for Middle Level Officers With Both Test and Rating Information*

| Dimension | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A Administrative | .. | | | | | | | | | | |
| B Communications | 50 | .. | | | | | | | | | |
| C Scheduling | 58 | 49 | .. | | | | | | | | |
| D Training | 51 | 51 | 45 | .. | | | | | | | |
| E Supervision | 48 | 53 | 50 | 43 | .. | | | D | | | |
| F Subordinate Evaluation | 51 | 57 | 42 | 60 | 44 | .. | | e | | | |
| G Field Command | 46 | 50 | 47 | 51 | 43 | 54 | .. | l | | | |
| H Public Relations | .. | .. | Deleted | | | .. | .. | e | | | |
| I Dedication | 37 | 38 | 44 | 39 | 41 | 41 | 45 | t | .. | | |
| J Overall Performance | 68 | 58 | 56 | 59 | 52 | 64 | 59 | e | 56 | .. | |
| K Mean of All Ratings | 78 | 76 | 75 | 75 | 73 | 77 | 75 | d | 65 | 85 | .. |

*Decimals omitted.

# APPENDIX D

# DECISION RULES USED FOR DEVELOPING MODAL RESPONSE KEYS FOR INVENTORIES OF LIKES AND DISLIKES AND SITUATIONAL JUDGMENT INVENTORIES

## SCORING RULES FOR DEVELOPING MODAL RESPONSE KEYS FOR THE LIKES AND DISLIKES INVENTORIES

The format has three separate sections: (a) the regular LID format; (b) the dichotomous format; and, (c) the blocks of ten items where respondent chooses the three he likes best and the three he likes least.

I. Here are the rules to be used for the format using blocks of ten:

1. First examine each *column* of ten percentages. Score the three items with the highest percentages under BEST as +1; score the three items with the highest percentages under LEAST as +1; score the four items with highest percentages under DIDN'T MARK as +1. In case of ties, weight both items.

2. Then examine response distributions for each row (i.e., each item) receiving scoring weights for BEST or LEAST. Give a weight of −1 to the response with the *smallest* percentage for each of these items. Do no further weighting for those items with weights for the DIDN'T MARK response.

II. Here are rules for the items with LID formats:

1. Examine L and D responses for each item. If either response has 60 percent or more, weight it +1; weight 1 as 0 and the other end of the item as −1.
Note that no Indifferent responses are to be weighted +1 on the basis of this rule alone.

2. Examine the sum of the percentages for 1 and the *larger* of the two percentages L or D. When the sum is 75 percent or greater, weight those two responses +1 and the third response −1. Do not weight items if the difference between L and D percentages is 5 percent or less.

III. Here are rules for items that are dichotomies:

1. If one response is 60 to 74 percent, score it +1, the other, 0.
2. If one response is 75 percent or greater, score it +1, the other, −1.

## SCORING RULES FOR DEVELOPING MODAL RESPONSE KEYS FOR SITUATIONAL JUDGMENT INVENTORIES

Here are decision rules for developing scoring keys for situational inventories for patrol officers, sergeants, detectives, and middle level command people:

1. The output gives the percentages of respondents who chose each response option as BEST for each item and the percentages who chose each response option as WORST. Subtracting the sum of percentages for BEST and WORST from 100 for each option gives the percentage who did not choose the option for either category. Use this "Not Used" choice as a basis for scoring along with the other choices of BEST and WORST.
2. First, examine the percentages for each response option for the choice BEST. Focus on the *two highest* percentages.
   a. If the two highest percentages differ by no more than 19 percentage points, weight each of the corresponding response options +1.
   b. If one percentage is 20 or more points larger than the other, weight the response option corresponding to the larger percentage +2 and the response option corresponding to the smaller percentage +1, *but if the smaller of the two percentages is less than 15, do not weight the corresponding response option at all.*
   c. Also, if the third highest percentage differs by four percentage points or less from the second highest percentage, then weight its corresponding response option +1 too.
3. Next, examine the percentages for each response option for the choice WORST. Again, focus on the *two highest* percentages, and fix scoring weights for response

options according to the decision rules delineated above for 2a, 2b, and 2c.

4. Now that response options have been weighted according to steps 2 and 3, give exactly opposite scoring weights to the response options for the opposite choice. For example:
   - If option 1 for a particular item had been given a weight of +2 for BEST, then give it a weight of −2 for WORST; similarly, if it had been given +1 for BEST, then give it a weight of −1 for WORST, etc.
   - If option 2 for a particular item had been given a weight of +2 for WORST, then give it a weight of −2 for BEST; similarly, if it had been given +1 for WORST, then give it a weight of −1 for BEST, etc.

   ***But note:***

   These rules occasionally lead to conflicting weights for a particular option for a particular choice. In these instances, remove all scoring weights from the "offending" response options. The content of such items seem to involve actions which may either be stated ambiguously or for which conventional wisdom could argue for either the BEST or WORST choice. They seem to represent "poor" response options, and it makes good sense not to weight them.)

5. Finally, examine the percentages for each response option for the "Not Used" choice. Don't weight any response option with 64 percent or fewer choices. Give +1 to response options with "Not Used" percentages between 65 percent and 84 percent. Give +2 to response options with "Not Used" percentages of 85 percent or greater.

# APPENDIX E

# TABLES SHOWING NUMBERS OF WEIGHTED ITEMS, MEDIAN FOLDBACK AND CROSS-VALIDITIES FOR VARIOUS SCORING KEYS ACCORDING TO ITEM POOL AND CRITERION SCORE LIMITS FOR EACH CRITERION DIMENSION AND EACH POLICE JOB FUNCTION

TABLE 1

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension P-Crime for Patrolmen*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV |
| Bio Past | 615 | 5 | .26 | .06 | 12 | .32 | .08 | 25 | .38 | .09 | 59 | .43 | .11 |
| Bio Mixed | 593 | 10 | .32 | .14 | 14 | .35 | .15 | 29 | .36 | .17 | 49 | .38 | .13 |
| Bio Total | 612 | 13 | .36 | .14 | 29 | .42 | .14 | 56 | .46 | .16 | 108 | .49 | .16 |
| Like-Dislike | 524 | 17 | .34 | .15 | 30 | .38 | .16 | 52 | .43 | .19 | — | — | — |
| Situational | 403 | 23 | .41 | .01 | 36 | .45 | .03 | 54 | .50 | −.05 | 95 | .54 | .01 |
| OSDI 1 | 524 | 1 | .12 | .05 | 4 | .21 | .06 | 15 | .31 | .10 | 41 | .35 | .11 |
| OSDI 2 | 529 | 0 | .00 | .11 | 2 | .23 | .10 | 9 | .33 | .13 | 31 | .35 | .15 |
| OSDI 3 | 529 | 5 | .29 | .17 | 13 | .32 | .15 | 30 | .34 | .16 | 57 | .32 | .17 |

TABLE 2

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension P-Force for Patrolmen*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV |
| Bio Past | 623 | 5 | .17 | .05 | 10 | .26 | .09 | 29 | .32 | .13 | 54 | .38 | .12 |
| Bio Mixed | 601 | 8 | .28 | .16 | 10 | .30 | .11 | 21 | .36 | .12 | 38 | .39 | .09 |
| Bio Total | 620 | 12 | .28 | .08 | 19 | .33 | .09 | 46 | .40 | .12 | 94 | .47 | .14 |
| Like-Dislike | 528 | 7 | .22 | .07 | 18 | .33 | .07 | 41 | .39 | .08 | — | — | — |
| Situational | 410 | 20 | .40 | .06 | 39 | .48 | .10 | 63 | .53 | .06 | 99 | .61 | .12 |
| OSDI 1 | 531 | 1 | .15 | .02 | 5 | .22 | .03 | 13 | .31 | .09 | 43 | .39 | .11 |
| OSDI 2 | 536 | 0 | .00 | .01 | 1 | .15 | .04 | 5 | .26 | .07 | 26 | .37 | .13 |
| OSDI 3 | 536 | 6 | .23 | .17 | 13 | .27 | .18 | 22 | .32 | .15 | 45 | .34 | .12 |

TABLE 3

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension P-Teamwork for Patrolmen*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV | No. of Items | Foldback | Mdn CV |
| Bio Past | 517 | 7 | .31 | .10 | 17 | .34 | .09 | 34 | .38 | .13 | 49 | .41 | .11 |
| Bio Mixed | 498 | 9 | .31 | .15 | 18 | .38 | .16 | 35 | .41 | .18 | 59 | .43 | .22 |
| Bio Total | 515 | 16 | .40 | .15 | 35 | .41 | .19 | 70 | .47 | .21 | 108 | .51 | .17 |
| Like-Dislike | 445 | 19 | .40 | .19 | 33 | .44 | .24 | 60 | .51 | .23 | — | — | — |
| Situational | 308 | 42 | .50 | −.00 | 58 | .53 | .05 | 80 | .60 | .04 | 115 | .61 | .03 |
| OSDI 1 | 441 | 7 | .28 | .15 | 10 | .32 | .14 | 20 | .39 | .14 | 51 | .46 | .13 |
| OSDI 2 | 443 | 0 | .00 | .05 | 4 | .24 | .09 | 16 | .36 | .12 | 42 | −.42 | .16 |
| OSDI 3 | 443 | 12 | .33 | .17 | 29 | .34 | .17 | 50 | .33 | .21 | 84 | .35 | .22 |

TABLE 4

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension P-Overall for Patrolmen*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 696 | 5 | .27 | .10 | 12 | .33 | .11 | 21 | .35 | .12 | 46 | .39 | .12 |
| Bio Mixed | 674 | 12 | .32 | .17 | 21 | .37 | .22 | 32 | .40 | .23 | 48 | .41 | .22 |
| Bio Total | 693 | 16 | .36 | .17 | 32 | .42 | .20 | 50 | .45 | .23 | 95 | .48 | .24 |
| Like-Dislike | 592 | 18 | .40 | .20 | 30 | .41 | .22 | 46 | .43 | .22 | 90 | .46 | .22 |
| Situational | 482 | 18 | .40 | .11 | 43 | .50 | .10 | 64 | .54 | .17 | 92 | .55 | .13 |
| OSDI 1 | 593 | 3 | .20 | −.00 | 6 | .22 | .15 | 18 | .35 | .18 | 44 | .39 | .23 |
| OSDI 2 | 600 | 2 | .23 | .15 | 2 | .23 | .17 | 21 | .30 | .21 | 47 | .37 | .23 |
| OSDI 3 | 600 | 10 | .30 | .20 | 21 | .33 | .22 | 43 | .29 | .19 | 68 | .29 | .23 |

TABLE 5

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension P-Random for Patrolmen*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 696 | 2 | .11 | −.02 | 4 | .15 | .04 | 18 | .31 | .07 | 37 | .38 | .04 |
| Bio Mixed | 674 | 1 | .09 | −.01 | 4 | .18 | .02 | 15 | .28 | .01 | 36 | .34 | .02 |
| Bio Total | 693 | 3 | .14 | −.01 | 7 | .19 | .01 | 30 | .37 | .02 | 69 | .44 | .05 |
| Like-Dislike | 592 | 3 | .17 | −.01 | 8 | .24 | .01 | 23 | .36 | .03 | — | — | — |
| Situational | 482 | 11 | .32 | −.03 | 23 | .39 | .00 | 49 | .46 | −.03 | 90 | .55 | .03 |
| OSDI 1 | 593 | 0 | .00 | −.04 | 1 | .08 | .04 | 6 | .20 | −.20 | 22 | .29 | −.04 |
| OSDI 2 | 600 | 0 | .00 | .00 | 0 | .00 | .08 | 3 | .19 | .04 | 9 | .26 | −.00 |
| OSDI 3 | 600 | 0 | .00 | .00 | 4 | .22 | .05 | 12 | .28 | .03 | 32 | .35 | .03 |

TABLE 6

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of the Item Pool, and Criterion Limits for Criterion Dimension D-Investigate for Detectives*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 377 | 16 | .38 | .03 | 20 | .41 | .09 | 38 | .48 | .05 | 65 | .52 | .04 |
| Bio Mixed | 363 | 13 | .38 | .09 | 22 | .45 | .08 | 37 | .50 | .12 | 64 | .56 | .14 |
| Bio Total | 374 | 29 | .45 | .11 | 38 | .51 | .09 | 69 | .57 | .20 | 131 | .61 | .18 |
| Like-Dislike | 304 | 20 | .38 | .10 | 40 | .47 | .11 | 69 | .50 | .12 | — | — | — |
| Situational | 359 | 29 | .46 | .10 | 49 | .55 | .06 | 79 | .62 | .07 | 116 | .63 | .11 |
| OSDI 1 | 218 | 11 | .36 | .65 | 21 | .44 | .06 | 40 | .53 | .05 | 62 | .57 | .05 |
| OSDI 2 | 296 | 0 | .00 | .00 | 3 | .23 | −.07 | 9 | .32 | −.01 | 28 | .41 | .00 |
| OSDI 3 | 296 | 8 | .27 | .07 | 16 | .34 | .03 | 30 | .42 | −.00 | 55 | .44 | .09 |

TABLE 7

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension D-Integrity for Detectives*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | Number of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 214 | 27 | .55 | .15 | 42 | .55 | .17 | 62 | .56 | .19 | 85 | .59 | .12 |
| Bio Mixed | 205 | 20 | .45 | −.04 | 32 | .56 | −.06 | 49 | .64 | .02 | 70 | .63 | .04 |
| Bio Total | 212 | 47 | .65 | .09 | 73 | .66 | .10 | 115 | .70 | .17 | 156 | .70 | .16 |
| Like-Dislike | 175 | 31 | .46 | −.08 | 53 | .57 | −.01 | 96 | .63 | .02 | — | — | — |
| Situational | 203 | 64 | .64 | .16 | 84 | .71 | .12 | 108 | .71 | .18 | 140 | .70 | .17 |
| OSDI 1 | 154 | 11 | .42 | .08 | 29 | .55 | .00 | 48 | .61 | .10 | 95 | .69 | .12 |
| OSDI 2 | 186 | 7 | .41 | .16 | 14 | .46 | .12 | 28 | .53 | .06 | 56 | .59 | .08 |
| OSDI 3 | 186 | 17 | .42 | .13 | 33 | .48 | .09 | 58 | .48 | .11 | 77 | .48 | .12 |

TABLE 8

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension D-Reporting for Detectives*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | Number of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 233 | 24 | .56 | .11 | 34 | .57 | .11 | 49 | .60 | .08 | 67 | .61 | .05 |
| Bio Mixed | 227 | 28 | .53 | .18 | 38 | .57 | .20 | 52 | .62 | .25 | 73 | .61 | .24 |
| Bio Total | 232 | 50 | .64 | .21 | 69 | .65 | .22 | 100 | .72 | .23 | 146 | .70 | .25 |
| Like-Dislike | 196 | 30 | .56 | .15 | 57 | .57 | .12 | 85 | .63 | .14 | — | — | — |
| Situational | 227 | 45 | .58 | .06 | 66 | .61 | .04 | 97 | .65 | .08 | 122 | .64 | .12 |
| OSDI 1 | 103 | 21 | .66 | −.06 | 39 | .68 | −.06 | 63 | .73 | −.02 | 88 | .75 | −.04 |
| OSDI 2 | 180 | 3 | .34 | .01 | 10 | .43 | .03 | 26 | .54 | −.01 | 57 | .53 | .03 |
| OSDI 3 | 180 | 22 | .52 | .04 | 29 | .56 | .17 | 51 | .57 | .08 | 75 | .57 | .12 |

TABLE 9

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension D-Overall for Detectives*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | Number of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 415 | 19 | .36 | .05 | 25 | .41 | .10 | 32 | .46 | .09 | 57 | .51 | .09 |
| Bio Mixed | 398 | 14 | .39 | .17 | 25 | .46 | .25 | 40 | .54 | .20 | 62 | .58 | .26 |
| Bio Total | 412 | 35 | .47 | .22 | 52 | .52 | .16 | 73 | .56 | .19 | 114 | .61 | .22 |
| Like-Dislike | 328 | 19 | .41 | .04 | 30 | .48 | .07 | 49 | .52 | .08 | — | — | — |
| Situational | 394 | 32 | .50 | .16 | 56 | .55 | .17 | 75 | .59 | .13 | 101 | .62 | .17 |
| OSDI 1 | 249 | 5 | .29 | .04 | 15 | .35 | −.01 | 36 | .43 | .02 | 61 | .52 | .04 |
| OSDI 2 | 329 | 0 | .00 | .00 | 2 | .21 | .06 | 8 | .32 | .02 | 25 | .39 | −.05 |
| OSDI 3 | 329 | 3 | .21 | .02 | 12 | .29 | .02 | 32 | .36 | .02 | 58 | .39 | .01 |

TABLE 10

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension D-Random for Detectives*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 — No. of Items | Foldback | Mdn CV | 48.0 and 52.5 — No. of Items | Foldback | Mdn CV | 48.5 and 51.5 — No. of Items | Foldback | Mdn CV | 49.0 and 51.0 — No. of Items | Foldback | Mdn CV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bio Past | 415 | 7 | .24 | −.00 | 17 | .31 | −.05 | 27 | .36 | −.02 | 52 | .44 | .01 |
| Bio Mixed | 398 | 8 | .28 | .00 | 13 | .35 | −.03 | 24 | .40 | .02 | 49 | .48 | −.01 |
| Bio Total | 412 | 16 | .38 | .04 | 32 | .43 | .02 | 53 | .48 | .00 | 102 | .57 | −.02 |
| Like-Dislike | 328 | 12 | .36 | −.03 | 20 | .38 | −.04 | 54 | .48 | −.03 | — | — | — |
| Situational | 394 | 15 | .34 | −.08 | 25 | .38 | −.11 | 55 | .48 | −.11 | 96 | .53 | −.09 |
| OSDI 1 | 249 | 12 | .37 | .13 | 20 | .38 | .08 | 31 | .44 | .05 | 61 | .49 | .06 |
| OSDI 2 | 329 | 0 | .00 | .00 | 0 | .00 | .00 | 5 | .27 | −.05 | 20 | .36 | −.05 |
| OSDI 3 | 329 | 2 | .19 | .03 | 10 | .33 | −.01 | 20 | .41 | .03 | 47 | .48 | −.02 |

TABLE 11

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension S-Training for Sergeants*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 — No. of Items | Foldback | Mdn CV | 48.0 and 52.0 — No. of Items | Foldback | Mdn CV | 48.5 and 51.5 — No. of Items | Foldback | Mdn CV | 49.0 and 51.0 — No. of Items | Foldback | Mdn CV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bio Past | 362 | 12 | .36 | .07 | 23 | .41 | .15 | 46 | .49 | .16 | 61 | .52 | .11 |
| Bio Mixed | 357 | 16 | .39 | .14 | 27 | .46 | .18 | 41 | .53 | .14 | 67 | .54 | .18 |
| Bio Total | 362 | 29 | .47 | .13 | 48 | .52 | .19 | 89 | .60 | .19 | 126 | .61 | .18 |
| Like-Dislike | 311 | 16 | .33 | .01 | 24 | .36 | .01 | 50 | .39 | −.04 | — | — | — |
| Situational | 353 | 40 | .55 | .10 | 57 | .60 | .08 | 77 | .60 | .07 | 104 | .60 | .11 |
| OSDI 1 | 280 | 10 | .43 | .09 | 13 | .43 | .08 | 24 | .51 | .02 | 54 | .54 | .05 |
| OSDI 2 | 338 | 1 | .17 | .07 | 5 | .33 | .10 | 16 | .38 | .12 | 38 | .42 | .09 |
| OSDI 3 | 338 | 13 | .32 | .10 | 17 | .31 | .09 | 34 | .33 | .09 | 51 | .35 | .05 |

TABLE 12

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension S-Supervision for Sergeants*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 — No. of Items | Foldback | Mdn CV | 48.0 and 52.0 — No. of Items | Foldback | Mdn CV | 48.5 and 51.5 — No. of Items | Foldback | Mdn CV | 49.0 and 51.0 — No. of Items | Foldback | Mdn CV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bio Past | 358 | 15 | .40 | .09 | 27 | .48 | .06 | 42 | .53 | .08 | 64 | .57 | .12 |
| Bio Mixed | 353 | 14 | .38 | .09 | 21 | .44 | .11 | 34 | .51 | .07 | 47 | .54 | .03 |
| Bio Total | 358 | 29 | .51 | .11 | 48 | .60 | .09 | 77 | .65 | .09 | 116 | .66 | .08 |
| Like-Dislike | 307 | 18 | .41 | .06 | 32 | .44 | .05 | 50 | .49 | −.01 | — | — | — |
| Situational | 349 | 30 | .50 | .09 | 54 | .56 | .06 | 78 | .61 | .13 | 112 | .64 | .10 |
| OSDI 1 | 278 | 10 | .41 | .08 | 19 | .48 | .07 | 28 | .52 | .08 | 52 | .57 | .07 |
| OSDI 2 | 335 | 0 | .00 | .03 | 4 | .25 | .07 | 14 | .39 | .06 | 30 | .45 | .09 |
| OSDI 3 | 335 | 5 | .29 | −.01 | 10 | .36 | −.03 | 22 | .37 | −.03 | 51 | .48 | −.02 |

TABLE 13

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension S-Inspection for Sergeants*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 283 | 16 | .43 | −.03 | 25 | .50 | .01 | 41 | .54 | .05 | 66 | .59 | .06 |
| Bio Mixed | 278 | 19 | .51 | .16 | 29 | .54 | .15 | 37 | .56 | .15 | 67 | .61 | .10 |
| Bio Total | 283 | 37 | .61 | .09 | 51 | .67 | .14 | 80 | .69 | .10 | 132 | .72 | .10 |
| Like-Dislike | 245 | 24 | .46 | .08 | 38 | .51 | .05 | 61 | .56 | .01 | — | — | — |
| Situational | 275 | 43 | .60 | .09 | 58 | .62 | .08 | 77 | .61 | .06 | 108 | .64 | .05 |
| OSDI 1 | 202 | 14 | .47 | .07 | 19 | .49 | .11 | 35 | .59 | −.01 | 72 | .65 | .06 |
| OSDI 2 | 259 | 4 | .29 | .09 | 6 | .34 | .01 | 15 | .43 | .08 | 35 | .49 | .04 |
| OSDI 3 | 259 | 7 | .29 | .06 | 17 | .38 | .02 | 31 | .43 | −.00 | 64 | .51 | .00 |

TABLE 14

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension S-Overall for Sergeants*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 361 | 13 | .39 | .10 | 19 | .45 | .08 | 40 | .52 | .12 | 68 | .55 | .10 |
| Bio Mixed | 356 | 13 | .42 | .13 | 22 | .47 | .12 | 39 | .52 | .09 | 62 | .52 | .08 |
| Bio Total | 361 | 28 | .53 | .14 | 39 | .58 | .17 | 76 | .63 | .13 | 129 | .64 | .12 |
| Like-Dislike | 310 | 10 | .32 | −.03 | 21 | .37 | −.05 | 38 | .40 | −.01 | — | — | — |
| Situational | 352 | 39 | .51 | .10 | 52 | .59 | .08 | 84 | .61 | .17 | 114 | .62 | .12 |
| OSDI 1 | 278 | 5 | .30 | −.00 | 14 | .41 | .05 | 35 | .56 | .09 | 63 | .59 | .10 |
| OSDI 2 | 336 | 0 | .00 | −.02 | 5 | .27 | .02 | 7 | .33 | .06 | 35 | .44 | .02 |
| OSDI 3 | 336 | 6 | .25 | .02 | 16 | .27 | .07 | 28 | .30 | .04 | 56 | .35 | .03 |

TABLE 15

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension S-Random for Sergeants*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 361 | 5 | .27 | −.02 | 13 | .36 | −.02 | 32 | .46 | .02 | 61 | .48 | −.02 |
| Bio Mixed | 356 | 12 | .34 | .00 | 22 | .42 | .07 | 42 | .49 | .04 | 64 | .55 | .09 |
| Bio Total | 361 | 18 | .41 | .06 | 35 | .52 | .05 | 74 | .61 | .07 | 126 | .65 | .01 |
| Like-Dislike | 310 | 8 | .28 | .01 | 21 | .38 | −.01 | 46 | .46 | −.04 | — | — | — |
| Situational | 352 | 18 | .41 | −.06 | 36 | .49 | −.02 | 65 | .58 | −.04 | 101 | .63 | −.02 |
| OSDI 1 | 278 | 8 | .37 | .11 | 14 | .45 | .09 | 35 | .48 | .14 | 60 | .51 | .13 |
| OSDI 2 | 336 | 0 | .00 | −.01 | 2 | .23 | .04 | 12 | .35 | .08 | 35 | .41 | .02 |
| OSDI 3 | 336 | 2 | .17 | −.08 | 5 | .24 | .00 | 26 | .31 | .00 | [illegible] | .32 | −.01 |

TABLE 16

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension C-Administration for Middle Command*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 — No. of Items | Foldback | Mdn CV | 48.0 and 52.0 — No. of Items | Foldback | Mdn CV | 48.5 and 51.5 — No. of Items | Foldback | Mdn CV | 49.0 and 51.0 — No. of Items | Foldback | Mdn CV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bio Past | 204 | 29 | .51 | .09 | 40 | .56 | .15 | 57 | .08 | .07 | 79 | .59 | .19 |
| Bio Mixed | 200 | 19 | .51 | .16 | 32 | .59 | .17 | 53 | .57 | .11 | 78 | .61 | .14 |
| Bio Total | 204 | 50 | .66 | .15 | 74 | .69 | .16 | 109 | .70 | .21 | 152 | .73 | .24 |
| Like-Dislike | 186 | 35 | .60 | .07 | 54 | .55 | .06 | 75 | .57 | .05 | — | — | — |
| Situational | 200 | 56 | .62 | .12 | 74 | .66 | .08 | 105 | .67 | .10 | 128 | .69 | .15 |
| OSDI 1 | 164 | 9 | .48 | −.04 | 18 | .55 | −.01 | 37 | .58 | −.03 | 72 | .64 | −.04 |
| OSDI 2 | 197 | 7 | .35 | .06 | 13 | .42 | .07 | 26 | .47 | .06 | 54 | .51 | .11 |
| OSDI 3 | 197 | 28 | .55 | .21 | 41 | .59 | .23 | 56 | .62 | .16 | 75 | .61 | .26 |

TABLE 17

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension C-Communications for Middle Command*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 — No. of Items | Foldback | Mdn CV | 48.0 and 52.0 — No. of Items | Foldback | Mdn CV | 48.5 and 51.5 — No. of Items | Foldback | Mdn CV | 49.0 and 51.0 — No. of Items | Foldback | Mdn CV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bio Past | 183 | 29 | .56 | .15 | 45 | .60 | .08 | 63 | .64 | .09 | 78 | .63 | .06 |
| Bio Mixed | 182 | 23 | .52 | .07 | 33 | .60 | .12 | 51 | .59 | .07 | 73 | .59 | .08 |
| Bio Total | 186 | 53 | .67 | .14 | 78 | .73 | .12 | 113 | .72 | .06 | 152 | .71 | .05 |
| Like-Dislike | 169 | 29 | .58 | .02 | 48 | .66 | .05 | 71 | .66 | .05 | — | — | — |
| Situational | 182 | 67 | .64 | .03 | 87 | .66 | .11 | 103 | .65 | .05 | 128 | .65 | .07 |
| OSDI 1 | 147 | 16 | .51 | .04 | 24 | .59 | .08 | 41 | .61 | .06 | 70 | .64 | .07 |
| OSDI 2 | 180 | 6 | .34 | −.03 | 14 | .43 | .07 | 33 | .48 | .01 | 63 | .51 | .07 |
| OSDI 3 | 180 | 20 | .52 | .17 | 34 | .54 | .09 | 55 | .56 | .09 | 75 | .55 | .11 |

TABLE 18

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension C-Personnel for Middle Command*

| Item Pool | Number of Persons | Criterion Limits for Non-Zero Scoring Weights: 47.5 and 52.5 — No. of Items | Foldback | Mdn CV | 48.0 and 52.0 — No. of Items | Foldback | Mdn CV | 48.5 and 51.5 — No. of Items | Foldback | Mdn CV | 49.0 and 51.0 — No. of Items | Foldback | Mdn CV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bio Past | 186 | 22 | .47 | −.03 | 40 | .57 | −.05 | 59 | .59 | .00 | 77 | .64 | −.03 |
| Bio Mixed | 182 | 30 | .58 | .10 | 39 | .59 | .10 | 54 | .59 | .11 | 74 | .63 | .19 |
| Bio Total | 186 | 48 | .66 | .09 | 79 | .71 | .08 | 110 | .71 | .11 | 151 | .72 | .09 |
| Like-Dislike | 169 | 29 | .53 | −.01 | 55 | .64 | −.01 | 86 | .68 | .01 | — | — | — |
| Situational | 182 | 65 | .68 | .10 | 84 | .68 | .17 | 109 | .73 | .11 | 133 | .74 | .14 |
| OSDI 1 | 147 | 25 | .59 | .05 | 39 | .62 | .11 | 65 | .70 | .11 | 91 | .71 | .10 |
| OSDI 2 | 180 | 6 | .31 | .03 | 16 | .42 | .01 | 28 | .44 | .04 | 55 | .53 | .01 |
| OSDI 3 | 180 | 23 | .53 | .08 | 34 | .54 | .13 | 49 | .57 | .12 | 68 | .55 | .12 |

TABLE 19

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension C-Dedication for Middle Command*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 157 | 31 | .62 | .10 | 44 | .65 | .09 | 57 | .68 | .11 | 78 | .68 | −.00 |
| Bio Mixed | 155 | 34 | .64 | .19 | 42 | .64 | .27 | 57 | .66 | .25 | 73 | .66 | .31 |
| Bio Total | 157 | 65 | .71 | .26 | 90 | .76 | .30 | 116 | .78 | .26 | 148 | .79 | .24 |
| Like-Dislike | 144 | 37 | .59 | −.03 | 54 | .67 | −.04 | 82 | .71 | .07 | — | — | — |
| Situational | 155 | 72 | .65 | .12 | 87 | .65 | .17 | 104 | .66 | .15 | 130 | .68 | .13 |
| OSDI 1 | 120 | 26 | .62 | .14 | 43 | .66 | .09 | 61 | .70 | .14 | 109 | .72 | .12 |
| OSDI 2 | 151 | 9 | .44 | −.03 | 14 | .51 | .06 | 36 | .61 | .00 | 62 | .64 | .03 |
| OSDI 3 | 151 | 19 | .48 | .00 | 33 | .46 | .03 | 47 | .50 | .00 | 65 | .53 | −.01 |

TABLE 20

*Number of Weighted Items, Foldback Coefficients, and Medi[illegible]ross-Validities for Scoring Keys, According to Type of Item Pool, and [illegible]erion Limits for Criterion Dimension C-Overall for Middle Command*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 206 | 28 | .57 | .13 | 41 | .61 | .15 | 57 | .62 | .14 | 83 | .64 | .16 |
| Bio Mixed | 202 | 26 | .51 | .13 | 39 | .56 | .22 | 59 | .58 | .24 | 79 | .61 | .20 |
| Bio Total | 206 | 50 | .67 | .16 | 75 | .70 | .26 | 114 | .69 | .30 | 161 | .71 | .34 |
| Like-Dislike | 188 | 41 | .58 | .16 | 59 | .54 | .17 | 78 | .60 | .18 | — | — | — |
| Situational | 202 | 58 | .60 | .19 | 77 | .61 | .13 | 104 | .67 | .18 | 123 | .69 | .23 |
| OSDI 1 | 166 | 14 | .47 | .02 | 23 | .53 | .00 | 48 | .56 | .06 | 80 | .62 | .06 |
| OSDI 2 | 199 | 7 | .36 | .05 | 16 | .46 | .12 | 33 | .50 | .18 | 62 | .52 | .07 |
| OSDI 3 | 199 | 28 | .54 | .21 | 37 | .56 | .16 | 49 | .57 | .12 | 71 | .58 | .17 |

TABLE 21

*Number of Weighted Items, Foldback Coefficients, and Median Cross-Validities for Scoring Keys, According to Type of Item Pool, and Criterion Limits for Criterion Dimension C-Random for Middle Command*

| | | Criterion Limits for Non-Zero Scoring Weights | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | 47.5 and 52.5 | | | 48.0 and 52.0 | | | 48.5 and 51.5 | | | 49.0 and 51.0 | | |
| Item Pool | of Persons | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV | No. of Items | Fold-back | Mdn CV |
| Bio Past | 206 | 30 | .54 | .07 | 40 | .55 | .05 | 52 | .54 | .02 | 71 | .58 | .06 |
| Bio Mixed | 202 | 19 | .54 | .06 | 28 | .58 | .02 | 43 | .63 | −.03 | 67 | .61 | −.02 |
| Bio Total | 206 | 50 | .65 | .11 | 69 | .68 | .04 | 95 | .69 | .03 | 136 | .72 | .04 |
| Like-Dislike | 188 | 25 | .55 | .06 | 42 | .59 | .05 | 71 | .66 | .04 | — | — | — |
| Situational | 202 | 43 | .63 | −.03 | 62 | .65 | −.07 | 84 | .69 | −.05 | 107 | .74 | −.02 |
| OSDI 1 | 166 | 9 | .36 | −.21 | 15 | .42 | −.22 | 32 | .54 | −.20 | 67 | .61 | −.22 |
| OSDI 2 | 199 | 3 | .26 | −.07 | 5 | .30 | −.10 | 18 | .44 | −.12 | 41 | .49 | −.11 |
| OSDI 3 | 199 | 7 | .29 | −.10 | 19 | .33 | −.14 | 36 | .47 | −.16 | 64 | .52 | −.11 |

# APPENDIX F

# TECHNICAL NOTE: PROCEDURE USED FOR ESTIMATING EXPECTED CROSS-VALIDITY COEFFICIENTS FOR FINAL PREDICTOR SCALES

Primarily for convenience in reporting, an approach was conceived for estimating the cross-validity of composite predictors formed by weighting two or more predictor variables when each was validated separately from the others. Although no claims of superior characteristics of the estimators are made, it is felt that they are based upon an acceptable rationale and lead to uninflated estimates.

Two types of problems presented themselves:

1. The predictor variables were separately validated with respect to the same criterion variable.
2. The predictor variables were each validated with respect to different (but correlated) criterion variables.

As shall be seen, the first of the two is a special case of the second.

*Definitions*

Let $p$ be a linear composite of $m$ predictor scores for any one person:

$$p = v_1x_1 + v_2x_2 + \ldots + v_mx_m$$

where $x_i$ ($i = 1, 2, \ldots, m$) is the "z-score" on the $i^{th}$ predictor and $v_i$ is the assigned weight. Then, let $z$ be a linear composite of n criterion scores for the same person:

$$z = w_1y_1 + w_2y_2 + \ldots + w_ny_n$$

where $y_i$ ($i = 1, 2, \ldots, n$) is a "z-score" for the $i^{th}$ criterion and $w_i$ is the assigned weight. By "z-score" it is meant that the score is expressed in terms of a transformation to a mean of zero and variance of one.

The value to be estimated is the "true" correlation between $p$ and $z$:

$$\rho_{pz} = \frac{\text{Cov}(p, z)}{[\text{Var}(p)\ \text{Var}(z)]^{\frac{1}{2}}}$$

where Cov (p, z) denotes the covariance of $p$ and $z$ and Var (x) denotes the variance of x.

It can be shown that:

$$\rho_{pz} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} v_iw_j\,\rho_{x_iy_j}}{\left[\left(\sum_{i=1}^{m}\sum_{j=1}^{m} v_iv_j\,\rho_{x_ix_j}\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_iw_j\,\rho_{y_iy_j}\right)\right]^{\frac{1}{2}}}$$

where $\rho_{x_iy_j}$ is the "true" correlation between the $i^{th}$ predictor and the $j^{th}$ criterion, $\rho_{x_ix_j}$, is the "true" intercorrelation between the $i^{th}$ and $j^{th}$ predictors, and $\rho_{y_iy_j}$ is the "true" intercorrelation between the $i^{th}$ and $j^{th}$ criteria.

*Estimation of cross-validity: single criterion variable*

Only sample statistics are available for the estimation of the cross-validities. For the case of a single criterion ($n = 1$), the following substitutions were made in the estimate of cross-validity, $\rho_{pz}$, after simplification allowed by the single criterion:

$$\hat{\rho}_{pz} = \frac{\sum_{i=1}^{m} v_i\,\hat{\rho}_{x_iy}}{\left(\sum_{i=1}^{m}\sum_{J=1}^{m} v_iv_j\,\hat{\rho}_{x_ix_j}\right)^{\frac{1}{2}}}$$

where $\hat{\rho}_{x_ix_j}$ is the Pearsonian coefficient computed between predictors in the develop-

mental sample and $\hat{\rho}_{x_i y}$ is the cross-validity estimate obtained in the Monte Carlo cross-validation procedure for the $i^{th}$ predictor with respect to the common criterion variable.

*Estimation of cross-validity: multiple criteria*

Again, the available sample statistics were substituted in order to obtain the single estimate of cross-validity for the case of n>1:

$$\hat{\rho}_{pc} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} v_i w_j \hat{\rho}_{x_i y_j}}{\left[\left(\sum_{i=1}^{m}\sum_{j=1}^{m} v_i v_j \hat{\rho}_{x_i x_j}\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \hat{\rho}_{y_i y_j}\right)\right]^{\frac{1}{2}}}$$

where, as before, $\hat{\rho}_{x_i x_j}$ is the Pearsonian coefficient between predictors in the developmental sample and $\hat{\rho}_{y_i y_j}$ is the Pearsonian coefficient between criterion variables in the sample.

The special problem of multiple criteria is found in the values to be used for $\hat{\rho}_{x_i y_j}$ because all combinations of predictors and criteria were not cross-validated through the Monte Carlo cross-validation procedure. In fact, any one predictor was cross-validated with respect to only one of the *n* criteria.

To fill in the missing values, the assumption of equal cross-validities was made so that the cross-validity obtained for a particular predictor was substituted as the estimate of validity for every criterion.

# APPENDIX G

# ITEM SAMPLES AND ILLUSTRATIVE SCORING WEIGHTS FOR CAREER INDEX PREDICTOR SCALES FOR PATROL OFFICERS DETECTIVES SERGEANTS MIDDLE LEVEL COMMAND OFFICERS

## PREDICTOR SCALE PI

| *Item* | *Positively Scored Response* |
|---|---|
| 1. I often act without thinking. | False |
| 2. It doesn't bother me to put aside what I have been doing without finishing it. | True |
| 3. I can honestly say that I do not mind paying my taxes because I feel that's one of the things I can do for what I get from the community. | True |
| 4. I daydream very little. | False |

## PREDICTOR SCALE PII

| *Item* | *Positively Scored Response* |
|---|---|
| 1. A person who doesn't vote is not a good citizen. | True |
| 2. In a group of people I would *not* be embarrassed to be called upon to start a discussion or give an opinion about something I know well. | True |
| 3. I have a tendency to give up easily when I meet difficult problems. | False |

## PREDICTOR SCALE PIII

| *Item* | *Positively Scored Response* |
|---|---|
| 1. Governor of a state | Like |
| 2. Detective stories | Like |
| 3. Inside work | Dislike |
| 4. Minding your own business in a conflict situation | Dislike |

## PREDICTOR SCALE PIV

*Item*

1. As a youngster, how often were you a leader in your group of friends?
   1. always
   2. frequently
   3. occasionally

   } POSITIVELY WEIGHTED

   4. seldom or never
   5. was not a member of a group
   6. can't remember

   } NEGATIVELY WEIGHTED
2. How sensitive have you been to criticism?
   1. much more sensitive than most
   2. more sensitive than most
   3. about as sensitive as most

   } POSITIVELY WEIGHTED

   4. less sensitive than most
   5. much less sensitive than most

   } NEGATIVELY WEIGHTED
3. Climbing along the edge of a precipice — "Like" is positively weighted
4. Directing traffic at a street corner — "Like" is positively weighted

## PREDICTOR SCALE DI

*Item*

1. What did you like best about high school?
   1. the chance to get a high school diploma ..... NEGATIVELY WEIGHTED

2. the activities in which you could participate
3. the friends you made
4. the things you learned

} POSITIVELY WEIGHTED

2. How important is it to you to work on a job where you can be friends with your co-workers and spend time talking to them?
   1. very little importance
   2. low importance

   } POSITIVELY WEIGHTED

   3. fairly important
   4. high importance

   } NOT WEIGHTED

   5. extreme importance .... NEGATIVELY WEIGHTED

PREDICTOR SCALE DII

| *Item* | *Positively Scored Response* |
|---|---|
| 1. I very much dislike making decisions when there is not enough available information. | True |
| 2. There is something wrong with a person who cannot take orders without getting angry and resentful. | True |
| 3. How many close relatives (parent, grandparent, uncle, brother) do you have who have been policemen? | |
| 1. none ..... NEGATIVELY WEIGHTED | |
| 2. 1–2<br>3. 3–5<br>4. 6 or more } POSITIVELY WEIGHTED | |

PREDICTOR SCALE SI

| *Item* | *Positively Scored Response* |
|---|---|
| 1. A patrolman on your shift has a drinking problem that is harming his performance, although as far as you know he has never come to work drunk or had a drink while on duty. You would: | |
| ____BEST 1. try to find out what his personal problems are to see why he is drinking. | |
| ____WORST 2. discuss it with his other supervisors. | |
| 3. warn him that if he ever seems at all under the influence, you will take disciplinary action. | Choose 5 as BEST |
| 4. let it go since you don't have any evidence that he is drunk on duty. | |
| 5. discuss it with your supervisor. | |
| 2. While off duty you are involved in a minor traffic accident. The patrolman at the scene is handling the situation incorrectly. You would: | |
| ____BEST 1. correct the patrolman. | |
| 2. speak to his sergeant about it the next day. | |
| ____WORST 3. correct the patrolman and also mention it to the patrolman's sergeant the next day. | Choose 3 as BEST |
| 4. ignore the situation because it was only a minor traffic accident. | |
| 5. say nothing about it to the patrolman's sergeant, but inform the captain. | |
| 3. During a spot inspection of a patrol car you find a private rifle one patrolman is carrying in the trunk. You would: | |
| ____BEST 1. confiscate the rifle. | |

____WORST 2. tell the officer to take it home when he goes off duty.
3. let it go since this officer always uses good judgment and might need the rifle on his patrol.
4. refer the incident to the lieutenant and record it in the patrolman's file.

Choose any answer but 1 as BEST

## PREDICTOR SCALE SII

*Item*

1. How much did you enjoy the work you did on any part-time jobs you have held?
   1. very much ..... POSITIVE
   2. much ..... NOT WEIGHTED
   3. somewhat ..... NOT WEIGHTED
   4. slightly ..... NEGATIVE
   5. not at all ..... NOT WEIGHTED
   6. never had a part-time job ..... NOT WEIGHTED

2. Using your own interpretation of what success means, do you feel you have been successful to this point in your life?
   1. yes
   2. partly

   } POSITIVELY WEIGHTED
   3. no
   4. not sure
   5. not old enough yet to say

   } NEGATIVELY WEIGHTED

3. The only reason a person works is to allow him/her to enjoy free time more.
   1. definitely agree
   2. probably agree

   } NEGATIVELY WEIGHTED
   3. not sure
   4. probably disagree
   5. definitely disagree

   } POSITIVELY WEIGHTED

## PREDICTOR SCALE CI

| *Item* | *Positively Scored Response* |
|---|---|
| 1. I have often lost out on things because I couldn't make up my mind soon enough. | False |
| 2. Many people will lie if it is to their advantage and they have the chance. | False |
| 3. How far did you go in school? | |
| 1. 8th grade or less<br>2. 9–11 years<br>3. graduated from high school | NEGATIVELY WEIGHTED |
| 4. 1–4 years of college<br>5. bachelor's degree | POSITIVELY WEIGHTED |
| 6. other ..... | NOT WEIGHTED |
| 4. How much did your parents emphasize success and getting ahead to you? | |
| 1. constantly ..... | POSITIVELY WEIGHTED |

2. frequently
3. now and then
4. rarely
5. never

NEGATIVELY WEIGHTED

## PREDICTOR SCALE CII

*Item*

1. When you were a high school student were you:
   1. one of the most active students
   2. more active than most students

   POSITIVELY WEIGHTED
   3. about as active as most students ..... NOT WEIGHTED
   4. not quite as active as most students ..... NEGATIVELY WEIGHTED
2. What has been your experience with people?
   1. there is a lot of good in all people ..... POSITIVELY WEIGHTED
   2. there is some good in most people ..... NOT WEIGHTED
   3. people are about as good as they have to be ... NEGATIVELY WEIGHTED
   4. a surprising number of people are mean and dishonest
   5. most people are just no good

   NOT WEIGHTED
3. When you were in high school, what positions did you hold?

*Positively Weighted Response*

. chairman of an important student committee
. editor of a publication
. member of the student council
. captain of an athletic team

## PREDICTOR SCALE CIII

*Item*

One of your best sergeants is typically 10 to 15 minutes late for work. You would:

____ BEST
____ WORST

1. dock his overtime a specified amount every day he is late.
2. keep a daily record, and report it in his personnel file.
3. ignore it because he is so competent the other seven hours and forty-five minutes.

| *Highest Positive Combination* | *Highest Negative Combination* |
|---|---|
| Choose 3 as WORST and either 1 or 2 as BEST | Choose 3 as BEST and either 1 or 2 as WORST |

An officer shows up for work smelling of liquor, but obviously not drunk. You would:

____ BEST
____ WORST

1. send him home and mark him "sick."
2. give him plenty of coffee, make sure he is sober, and send him on patrol.
3. reassign him for the day to station duty.
4. send him home and require that he make it up on one of his off days.

| *Highest Positive Combination* | *Highest Negative Combination* |
|---|---|
| Choose 4 as BEST and 2 as WORST | Choose 2 as BEST and 4 as WORST |

## PREDICTOR SCALE CIV

*Item*

1. On the average, how many hours a week did you work on a part-time job in high school?
   1. none ..... NEGATIVELY WEIGHTED
   2. 1–4 ..... NOT WEIGHTED
   3. 5–10 ..... NOT WEIGHTED
   4. 11–15 ..... POSITIVELY WEIGHTED
   5. 16 or more ..... NOT WEIGHTED

2. Some people are completely involved in their job—they are absorbed in it night and day. For other people, their job is simply one of several interests. How involved do you usually feel in a job?
   1. very little involved—my other interests are more absorbing } NOT WEIGHTED
   2. slightly involved
   3. moderately involved—my job and other interests are equally absorbing to me } NEGATIVELY WEIGHTED
   4. strongly involved
   5. very strongly involved—my work is the most absorbing interest in my life } POSITIVELY WEIGHTED

## PREDICTOR SCALE CV

| *Item* | *Positively Scored Response* |
|---|---|
| 1. You, as lieutenant, want to assess the training needs of your patrolmen. You would: | |
| ___ BEST 1. ask the patrolmen what they need to know better. | |
| ___ WORST 2. ask both the patrolmen and the sergeants. | |
| 3. look over the reports and performance reviews of the men to pinpoint their weaknesses. | Choose 4 as WORST |
| 4. plan a general review of the academy course, since all the men could use a refresher. | |
| 5. consult with the judges and district attorney to learn how to improve your conviction rate. | |
| 2. A call comes in for assistance at the scene of a gang fight in a ghetto area. The report says that one gang is predominantly Chicano and has about 15 members, and the other gang is predominantly black and has about 20 members. You, the captain, would: | |
| ___ BEST 1. dispatch 3 squads to the area. | Choose 1 or 3 as BEST |
| 2. dispatch 2 squads to the scene, plus the | |
| ___ WORST lieutenant, because he needs experience in such situations. | If 3 is chosen as WORST, it is *negatively* weighted |
| 3. dispatch two squads and rush to the scene yourself to coordinate the efforts. | |
| 4. dispatch one squad, the lieutenant, and yourself to the scene. | |
| 5. send the sergeant and all available precinct squads. | |

# APPENDIX H

# VALIDITIES FOR PREDICTOR SCALES AGAINST CRITERION DIMENSIONS FOR TOTAL SAMPLES AND SEPARATELY BY CITY

TABLE 1

*Validities* for Predictor Scales of the Patrol Officer Career Index for Total Sample and Separately by City** (Whites Only)*

| Predictor Scale | Criterion Dimension: Crime | Force | Teamwork | Overall | Random |
|---|---|---|---|---|---|
| *PI* | | | | | |
| Total (502) | 18 | 40 | 13 | 24 | 01 |
| Minneapolis (155) | 17 | 43 | 17 | 25 | −05 |
| Portland (47–49) | 27 | 46 | 06 | 28 | −33 |
| Des Moines (87–91) | 00 | 39 | 05 | 11 | 06 |
| Cincinnati (92–93) | 20 | 33 | 03 | 17 | 02 |
| Miami (50) | 22 | 00 | 01 | 03 | 22 |
| Washington, D. C. (60–62) | 20 | 46 | −09 | 32 | −04 |
| *PII* | | | | | |
| Total (452) | 29 | 14 | 32 | 37 | 03 |
| Minneapolis (153) | 26 | 18 | 28 | 32 | −10 |
| Portland (46–48) | 43 | 35 | 22 | 41 | −19 |
| Des Moines (84–88) | 09 | 18 | 08 | 10 | −01 |
| Cincinnati (92–93) | 24 | 11 | 06 | 22 | 06 |
| Miami (49) | 18 | −11 | 16 | 21 | 18 |
| Washington, D. C. (60–62) | 36 | 32 | 28 | 49 | −16 |
| *PIII* | | | | | |
| Total (528) | 41 | 11 | 28 | 42 | 04 |
| Minneapolis (152) | 42 | −01 | 15 | 37 | 11 |
| Portland (49–51) | 40 | 17 | 20 | 24 | −09 |
| Des Moines (72–74) | 39 | 17 | 07 | 32 | 02 |
| Cincinnati (89–90) | 40 | 14 | 23 | 38 | 03 |
| Miami (52) | 25 | −14 | 09 | 34 | 22 |
| Washington, D. C. (55–57) | 03 | −19 | 51 | 24 | −06 |
| *PIV* | | | | | |
| Total (372) | 37 | 06 | 54 | 48 | 03 |
| Minneapolis (139) | 38 | 03 | 48 | 43 | −06 |
| Portland (46–47) | 63 | 27 | 52 | 43 | −05 |
| Des Moines (72–74) | 22 | 07 | 09 | 24 | 06 |
| Cincinnati (91–92) | 38 | 18 | 43 | 43 | 02 |
| Miami (53) | 38 | −02 | 26 | 35 | 12 |
| Washington, D. C. (56–58) | 03 | −19 | 51 | 24 | −06 |

*Decimals omitted.

**Ns vary according to criterion dimension. The ranges of sample sizes are shown within the parentheses.

TABLE 2

*Validities* for Predictor Scales of the Detective Career Index for Total Sample and Separately by City** (Whites Only)*

| Predictor Scale | Criterion Dimension: Investigation | Integrity | Reporting | Overall | Random |
|---|---|---|---|---|---|
| *DI* | | | | | |
| Total (245) | 52 | 23 | 58 | 59 | 00 |
| Minneapolis (44) | 69 | 46 | 61 | 76 | −02 |
| Portland (42) | 41 | −03 | 55 | 57 | −07 |
| Cincinnati (39–41) | 66 | 40 | 53 | 65 | 09 |
| Washington, D. C. (36–42) | 37 | 06 | 47 | 64 | 12 |
| San Diego (52) | 41 | 13 | 40 | 42 | 01 |
| Des Moines, Tucson, Albuquerque (49) | 58 | 50 | 61 | 70 | −03 |
| *DII* | | | | | |
| Total (263) | 11 | 62 | 26 | 17 | −09 |
| Minneapolis (40) | 13 | 50 | 26 | 21 | −21 |
| Portland (42) | 03 | 75 | 14 | 35 | −27 |
| Cincinnati (20–21) | −15 | 12 | 28 | −16 | 44 |
| Washington, D. C. (39–45) | 31 | 49 | 03 | 19 | −13 |
| San Diego (39) | 03 | −11 | 27 | 05 | −14 |
| Des Moines, Tucson, Albuquerque (52) | 06 | 42 | 12 | 08 | 00 |

*Decimals omitted.

**Ns vary according to criterion dimension. The ranges of sample sizes are shown within the parentheses.

TABLE 3

*Validities* for Predictor Scales of the Sergeant Career Index for Total Sample and Separately by City** (Whites Only)*

| *Predictor Scale* | *Criterion Dimension* | | | | |
|---|---|---|---|---|---|
| | *Training* | *Supervision* | *Inspection* | *Overall* | *Random* |
| *SI* | | | | | |
| Total (271) | 45 | 55 | 48 | 59 | 00 |
| Minneapolis (38–39) | 56 | 62 | 47 | 64 | −01 |
| Des Moines (37–39) | 28 | 52 | 31 | 55 | 20 |
| Cincinnati (61) | 36 | 35 | 25 | 47 | −09 |
| Miami (42–44) | 50 | 61 | 58 | 68 | −25 |
| Washington, D. C. (55) | 38 | 50 | 37 | 58 | 21 |
| San Diego (41–42) | 56 | 45 | 56 | 54 | 03 |
| *SII* | | | | | |
| Total (286) | 54 | 37 | 40 | 54 | 06 |
| Minneapolis (37–38) | 53 | 28 | 45 | 48 | −24 |
| Des Moines (31–33) | 36 | 20 | 42 | 48 | 32 |
| Cincinnati (63–64) | 41 | 41 | 28 | 51 | −04 |
| Miami (42–45) | 73 | 54 | 57 | 62 | 06 |
| Washington, D. C. (57) | 58 | 43 | 37 | 62 | 30 |
| San Diego (42) | 44 | 35 | 31 | 52 | −08 |

*Decimals omitted.

**Ns vary according to criterion dimension. The ranges of sample sizes are shown within the parentheses.

TABLE 4

*Validities* for Predictor Scales of the Middle Command Career Index for Total Sample and Separately by City** (Whites Only)*

| Predictor Scale | Criterion Dimension: Administration | Communications | Personnel | Dedication | Overall | Random |
|---|---|---|---|---|---|---|
| *CI* | | | | | | |
| Total (123) | 76 | 42 | 43 | 35 | 65 | 07 |
| Minneapolis (27) | 63 | 14 | 25 | 23 | 38 | 15 |
| Portland (17) | 78 | 33 | 59 | 25 | 59 | −46 |
| Des Moines (19) | 89 | 67 | 53 | 48 | 83 | −16 |
| Cincinnati (22–24) | 81 | 78 | 62 | 52 | 78 | 30 |
| Miami (28) | 69 | 53 | 59 | 61 | 77 | 17 |
| Washington, D. C. (41) | 72 | 30 | 21 | 44 | 56 | 12 |
| San Diego (27–28) | 82 | 31 | 43 | 05 | 46 | 22 |
| *CII* | | | | | | |
| Total (140) | 55 | 51 | 52 | 48 | 72 | 01 |
| Minneapolis (28) | 35 | 39 | 34 | 30 | 50 | 05 |
| Portland (18) | 65 | 48 | 51 | 35 | 74 | −40 |
| Des Moines (19) | 68 | 58 | 42 | 52 | 70 | −16 |
| Cincinnati (22–24) | 61 | 76 | 71 | 62 | 79 | −13 |
| Miami (28) | 40 | 53 | 61 | 67 | 72 | 02 |
| Washington, D. C. (43) | 62 | 29 | 32 | 39 | 64 | 23 |
| San Diego (30) | 51 | 60 | 52 | 25 | 71 | 06 |
| *CIII* | | | | | | |
| Total (193) | 40 | 32 | 36 | 26 | 48 | 05 |
| Minneapolis (25) | −28 | 04 | −12 | 04 | −04 | 12 |
| Portland (18) | 72 | 40 | 59 | 40 | 62 | −43 |
| Des Moines (20) | 60 | 56 | 48 | 57 | 74 | 12 |
| Cincinnati (21–23) | 56 | 54 | 51 | 55 | 60 | −04 |
| Miami (29) | 38 | 17 | 38 | 06 | 29 | 09 |
| Washington, D. C. (39) | 45 | 29 | 63 | 29 | 57 | 26 |
| San Diego (30) | 19 | 37 | 41 | −08 | 45 | −13 |
| *CIV* | | | | | | |
| Total (158) | 27 | 27 | 31 | 77 | 44 | −08 |
| Minneapolis (28) | 42 | 27 | 32 | 74 | 61 | −17 |
| Portland (18) | 56 | 40 | 46 | 74 | 70 | −25 |
| Des Moines (19) | 21 | 47 | 19 | 87 | 53 | −05 |
| Cincinnati (22–24) | 39 | 32 | 41 | 73 | 51 | 11 |
| Miami (28) | 17 | 57 | 41 | 82 | 44 | −14 |
| Washington, D. C. (42) | 08 | −14 | 10 | 24 | −03 | −15 |
| San Diego (30) | 23 | 18 | 14 | 77 | 39 | −15 |
| *CV* | | | | | | |
| Total (131) | 37 | 50 | 74 | 38 | 61 | −08 |
| Minneapolis (25) | 32 | 41 | 62 | 02 | 55 | 00 |
| Portland (17) | 59 | 70 | 59 | 29 | 66 | −50 |

| Predictor Scale | Criterion Dimension | | | | | |
|---|---|---|---|---|---|---|
| | Administration | Communications | Personnel | Dedication | Overall | Random |
| Des Moines (19) | 08 | 21 | −18 | 40 | 20 | 00 |
| Cincinnati (20–22) | 76 | 69 | 88 | 71 | 86 | −08 |
| Miami (27) | 43 | 50 | 78 | 73 | 75 | −08 |
| Washington, D. C. (40) | 33 | 31 | 62 | 35 | 46 | −04 |
| San Diego (30) | 41 | 49 | 80 | 03 | 68 | 04 |

*Decimals omitted.

**Ns vary according to criterion dimension. The ranges of sample sizes are shown within the parentheses.

☆U.S. GOVERNMENT PRINTING OFFICE: 1976 O—204-656